

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number  
**WO 02/29086 A2**

(51) International Patent Classification<sup>7</sup>: **C12Q**

(US). **LEWIS, Marcia, E.** [US/US]; 67 Wheelwright Farm, Cohasset, MA 02025 (US).

(21) International Application Number: PCT/US01/30732

(22) International Filing Date: 2 October 2001 (02.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/237,271 2 October 2000 (02.10.2000) US

(71) Applicant (*for all designated States except US*): **BAYER CORPORATION** [US/US]; 33 Coney Street, East Walpole, MA 02032 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **BURGESS, Christopher** [US/US]; 97 Canton Terrace, Westwood, MA 02090 (US). **ASTLE, John, H.** [US/US]; 42 Short Sreet, Taunton, MA 02780 (US). **CARROLL, Eddie, III** [US/US]; 1175 Washington Street, Norwood, MA 02062 (US). **CATINO, Theodore, J.** [US/US]; 18 Jo Paul Drive, Attleboro, MA 02702 (US). **DWIVEDI, Poornima** [US/US]; 10 Haven Road, Medfield, MA 02052 (US). **MOLINO, Gary, A.** [US/US]; 3 Essex Street, Norfolk, MA 02056 (US). **THIAGLINGAM, Arunthathi** [US/US]; 26 Winchestrer Drive, Lexington, MA 02420

(74) Agent: **EVANS, Paula, Campbell**; Palmer & Dodge, LLP, One Beacon Street, Boston, MA 02108 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: NUCLEIC ACID SEQUENCES DIFFERENTIALLY EXPRESSED IN CANCER TISSUE

(57) Abstract: This invention relates to novel nucleic acid sequences which are differentially expressed in cancer cells. The invention also relates to proteins and peptides encoded by the sequences, to diagnostic assays and therapeutic agents based on the sequences and proteins, and to probes, antisense constructs, and antibodies derived from the sequences and proteins or peptides. The subject nucleic acids have been found to be differentially expressed by tumor cells, particularly in colon cancer tissue.



WO 02/29086 A2

NUCLEIC ACID SEQUENCES DIFFERENTIALLY EXPRESSED IN CANCER TISSUEField of the Invention

5 The present invention provides nucleic acid sequences and proteins encoded thereby which are differentially expressed in cancer tissues, as well as probes derived from the nucleic acid sequences, antibodies directed to the encoded proteins, and diagnostic methods for determining the presence and state of cancerous cells, especially colon cancer cells.

Background of the Invention

10 Colorectal carcinoma is a malignant neoplastic disease. There is a high incidence of colorectal carcinoma in the Western world, particularly in the United States. Tumors of this type often metastasize through lymphatic and vascular channels. Many patients with colorectal carcinoma eventually die from this disease. In fact, it is estimated that 62,000 persons in the United States alone die of colorectal carcinoma annually.

15 However, if diagnosed early, colon cancer may be treated effectively by surgical removal of the cancerous tissue. Colorectal cancers originate in the colorectal epithelium and typically are not extensively vascularized (and therefore not invasive) during the early stages of development. Colorectal cancer is thought to result from the clonal expansion of a single mutant cell in the epithelial lining of the colon or rectum. The transition to a highly vascularized, invasive and ultimately metastatic cancer which spreads throughout the body commonly takes  
20 ten years or longer. If the cancer is detected prior to invasion, surgical removal of the cancerous tissue is an effective cure. However, colorectal cancer is often detected only upon manifestation of clinical symptoms, such as pain and black tarry stool. Generally, such symptoms are present only when the disease is well established, often after metastasis has occurred, and the prognosis for the patient is poor, even after surgical resection of the cancerous tissue. Early detection of  
25 colorectal cancer therefore is important in that detection may significantly reduce its morbidity.

Invasive diagnostic methods such as endoscopic examination allow for direct visual identification, removal, and biopsy of potentially cancerous growths such as polyps. Endoscopy is expensive, uncomfortable, inherently risky, and therefore not a practical tool for screening populations to identify those with colorectal cancer. Non-invasive analysis of stool samples for  
30 characteristics indicative of the presence of colorectal cancer or precancer is a preferred

alternative for early diagnosis, but no known diagnostic method is available which reliably achieves this goal.

### Summary of the Invention

5 The present invention provides nucleic acid sequences and proteins encoded thereby, as well as probes derived from the nucleic acid sequences, antibodies directed to the encoded proteins, and diagnostic methods for detecting cancerous cells, especially colon cancer cells. The sequences disclosed herein have been found to be differentially expressed in colon cancer cell lines and/or colon cancer tissue.

10 In one aspect, the invention provides an isolated nucleic acid sequence comprising SEQ ID Nos 1-503, or a sequence complementary thereto.

In another aspect, the invention provides an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto.

15 In another embodiment, the nucleic acid is at least about 80% to about 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto.

20 In another aspect, the invention provides an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80% or about 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503 or a sequence complementary thereto.

25 In one embodiment, the invention provides a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, and a transcriptional regulatory sequence operably linked to the nucleotide sequence to render the nucleotide sequence suitable for use as an expression vector. In another embodiment, the nucleic acid may be

included in an expression vector capable of replicating in a prokaryotic or eukaryotic cell. In a related embodiment, the invention provides a host cell transfected with the expression vector.

5 In another embodiment, the invention provides a transgenic animal having a transgene of a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-1103, preferably SEQ ID Nos 1-503, or a sequence complementary thereto incorporated in cells thereof. The transgene modifies the level of expression of the nucleic acid, the stability of a mRNA transcript of the nucleic acid, or the activity of the encoded product of the nucleic acid.

10 In yet another embodiment, the invention provides a substantially pure nucleic acid comprising the nucleotide sequence of SEQ ID Nos 1-1103, or a sequence complementary thereto.

In yet another embodiment, the invention provides a substantially pure nucleic acid which hybridizes under stringent conditions to a nucleic acid probe corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos 1-503, or a sequence complementary thereto.

The invention also provides an antisense oligonucleotide analog which hybridizes under stringent conditions to at least 12, at least 25, or at least 50 consecutive nucleotides of one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 up to the full length of one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto, and which is resistant to cleavage by a nuclease, preferably an endogenous endonuclease or exonuclease.

25 In another embodiment, the invention provides a probe/primer comprising a substantially purified oligonucleotide comprising at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of SEQ ID Nos 1-1103, or a sequence complementary thereto.

In another embodiment, the invention provides a probe/primer comprising a substantially purified oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos. 1-1103 up to the full length of one of SEQ ID Nos. 1-1103 or a sequence complementary thereto. In preferred embodiments, the probe selectively hybridizes with a target nucleic acid. In



another embodiment, the probe may include a label group attached thereto and able to be detected. The label group may be selected from radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors. The invention further provides arrays of at least about 10, at least about 25, at least about 50, or at least about 100 different probes as described above attached to a solid support.

In yet another embodiment, the invention pertains to a method of determining the phenotype of a cell comprising detecting the differential expression, relative to a normal cell, of at least one nucleic acid of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, wherein the nucleic acid is differentially expressed by at least a factor of two, at least a factor of five, at least a factor of twenty, or at least a factor of fifty.

In a still further embodiment, the invention pertains to a method of determining the phenotype of cell, comprising detecting the differential expression, relative to a normal cell, of at least one protein encoded by a nucleic acid which hybridizes under stringent conditions to a sequence selected from the group consisting of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, wherein the protein is differentially expressed by at least a factor of two, at least a factor of five, at least a factor of twenty, an up to at least a factor of 50.

The invention further provides a method of determining the phenotype of cell, comprising detecting the differential expression, relative to a normal cell, of at least one polypeptide selected from the group of polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, wherein the polypeptide is differentially expressed by at least a factor of two, at least a factor of five, at least a factor of twenty, an up to at least a factor of 50.

In yet another embodiment, the invention pertains to a method of determining the phenotype of a cell comprising detecting the differential expression, relative to a normal cell, of at least one nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, wherein the nucleic acid is differentially expressed by at least a factor of two, at least a factor of five, at least a factor of twenty, or at least a factor of fifty.

In another aspect, the invention provides polypeptides encoded by the subject nucleic acids. In one embodiment, the invention pertains to a polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-1103 or a sequence complementary thereto,

or a fragment comprising at least about 25, or at least about 40 amino acids thereof. Further provided are antibodies immunoreactive with these polypeptides.

In a further aspect the invention pertains to a polypeptide encoded by one or more of the sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494.

In a still further aspect the invention pertains to a polypeptide having the sequence of one or SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 44857, 4489, 4491, and 4493.

In still another aspect, the invention provides diagnostic methods. In one embodiment, the invention pertains to a method for determining the phenotype of cells from a patient by providing a nucleic acid probe comprising a nucleotide sequence having at least 10, at least about 15, at least about 25, or at least about 40 consecutive nucleotides represented in a sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 up to the full length of one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto, obtaining a sample of cells from a patient, optionally providing a second sample of cells substantially all of which are non-cancerous, contacting the nucleic acid probe under stringent conditions with mRNA of each of said first and second cell samples, and comparing (a) the amount of hybridization of the probe with mRNA of the first cell sample, with (b) the amount of hybridization of the probe with mRNA of the second cell sample, wherein a difference of at least a factor of two, at least a factor of five, at least a factor of twenty, or at least a factor of fifty in the amount of hybridization with the mRNA of the first cell sample as compared to the amount of hybridization with the mRNA of the second cell sample is indicative of the phenotype of cells in the first cell sample. Determining the phenotype includes determining the genotype, as the term is used herein.

In another embodiment, the invention provides a test kit for identifying the presence of cancerous cells or tissues, comprising a probe/primer as described above, for measuring a level of a nucleic acid which hybridizes under stringent conditions to a nucleic acid of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 in a sample of cells isolated from a patient. In certain embodiments, the kit may further include instructions for using the kit, solutions for suspending or fixing the cells, detectable tags or labels, solutions for rendering a nucleic acid susceptible to hybridization, solutions for lysing cells, or solutions for the purification of nucleic acids.

In another embodiment, the invention provides a method of determining the phenotype of a cell, comprising detecting the differential expression, relative to a normal or control cell, of at least one protein encoded by a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, wherein the protein is differentially expressed by at least a factor of two, at least a factor of five, at least a factor of twenty, or at least a factor of fifty. In one embodiment, the level of the protein is detected in an immunoassay. The invention also pertains to a method for determining the presence or absence of a nucleic acid, such as mRNA, which hybridizes under stringent conditions to one of SEQ ID Nos. 1-1103 in a cell, comprising contacting the cell with a probe as described above. The invention further provides a method for determining the presence or absence of a subject polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-1103 in a cell, comprising contacting the cell with an antibody as described above.

In yet another embodiment, the invention provides a method for determining the presence of an aberrant mutation (e.g., deletion, insertion, or substitution of nucleic acids) or aberrant methylation in a sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-1103 or a sequence complementary thereto, comprising collecting a sample of cells from a patient, isolating nucleic acid from the cells of the sample, contacting the nucleic acid sample with one or more probe/primers which specifically hybridize to a nucleic acid sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, under conditions such that hybridization and/or amplification of the nucleic acid occurs, and comparing the presence, absence, or size of an amplification product to the amplification product of a normal cell.

In one embodiment, the invention provides a test kit for identifying the presence of cancer cells, comprising an antibody specific for a protein encoded by a nucleic acid which hybridizes under stringent conditions to any one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto. In certain embodiments, the kit further includes instructions for using the kit. In certain embodiments, the kit may further include solutions for suspending or fixing the cells, detectable tags or labels, solutions for rendering a polypeptide susceptible to the binding of an antibody, solutions for lysing cells, or solutions for the purification of polypeptides.

In yet another aspect, the invention provides pharmaceutical compositions including the subject nucleic acids. In one embodiment, an agent which alters the level of expression in a cell

of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto is identified by providing a cell, treating the cell with a test agent, determining the level of expression in the cell of a nucleic acid which hybridizes under stringent  
5 conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto, and comparing the level of expression of the nucleic acid in the treated cell with the level of expression of the nucleic acid in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of expression of the nucleic acid in the untreated cell is indicative of an  
10 agent which alters the level of expression of the nucleic acid in a cell. The invention further provides a pharmaceutical composition comprising an agent identified by this method. In another embodiment, the invention provides a pharmaceutical composition which includes a polypeptide encoded by a nucleic acid having a nucleotide sequence that hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484,  
15 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto. In one embodiment, the invention pertains to a pharmaceutical composition comprising a nucleic acid including a sequence which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto.

20 In yet another aspect, the invention provides pharmaceutical compositions including the subject nucleic acids. In one embodiment, an agent which alters the level of expression in a cell of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto is identified by providing a cell, treating the cell with a test agent, determining the level  
25 of expression in the cell of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto, and comparing the level of expression of the nucleic acid in the treated cell with the level of expression of the nucleic acid in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of  
30 expression of the nucleic acid in the untreated cell is indicative of an agent which alters the level of expression of the nucleic acid in a cell.

The invention further provides a method for identifying an agent which alters the level of expression in a cell of a polypeptide having a sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493 comprising providing a cell; treating the

cell with the test agent; determining the level of expression of one or more polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493 in the cell by reacting the cell with an antibody specific for one or more of the polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493; and  
5 comparing the level of expression of the polypeptide in the treated cell with the level of expression of the same polypeptide in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of expression of the nucleic acid in the untreated cell is indicative of an agent which alters the level of expression of the polypeptide in a cell.

10 The invention further provides a pharmaceutical composition comprising an agent identified by the above methods. In another embodiment, the invention provides a pharmaceutical composition which includes a polypeptide encoded by a nucleic acid having a nucleotide sequence that hybridizes under stringent conditions to one of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence  
15 complementary thereto. In a further embodiment the invention provides a pharmaceutical composition comprising one or more antibodies which bind to a polypeptide encoded by one or more of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. In a still further embodiment, the invention provides a pharmaceutical composition comprising one or more antibodies which binds to a polypeptide of one or more of SEQ ID Nos.  
20 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. In one embodiment, the invention pertains to a pharmaceutical composition comprising a nucleic acid including a sequence which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or a sequence complementary thereto.

25 In one embodiment the invention relates to a method for detecting cancer in a patient sample in which an antibody to a protein encoded by SEQ ID Nos 1-4470 is used to react with proteins in the patient sample. In a further embodiment, the invention relates to a method for detecting cancer in a patient sample in which an antibody to a protein encoded by one or more of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 is  
30 used to react with proteins in the patient sample. In a still further embodiment, the invention provides a method for detecting cancer in a patient sample in which an antibody to a protein having the sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493 is used to react with protein in the patient sample.

### Brief Description of the Figure

Figure 1 depicts the nucleic acid sequence of SEQ ID Nos: 1-4470.

Figure 2 depicts the nucleic acid sequence of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494.

5        Figure 3 depicts the amino acid sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493.

### Detailed Description of the Invention

The invention relates to nucleic acids having the disclosed nucleotide sequences (SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494), as  
10    well as full length cDNA, mRNA, and genes corresponding to these sequences, and to polypeptides and proteins encoded by these nucleic acids and genes, and portions thereof. In particular the invention relates to the full length cDNA sequence of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 and the polypeptide sequence encoded thereby and shown in SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485,  
15    4487, 4489, 4491, and 4493, respectively. The 4494 sequences disclosed herein were analyzed by comparing the sequences to those disclosed in publicly available databases. Based upon the search results, it was found that SEQ ID Nos: 1-503 contained novel sequences, SEQ ID Nos: 504-1103 contained known EST sequences, and SEQ ID Nos: 1104-4494 contained known sequences.

20        Also included in the present invention are polypeptides and proteins encoded by the nucleic acids of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, and in particular the polypeptide sequences of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. The various nucleic acids that can encode these polypeptides and proteins differ because of the degeneracy of the genetic code,  
25    in that most amino acids are encoded by more than one triplet codon. The identity of such codons is well known in this art, and this information can be used for the construction of the nucleic acids within the scope of the invention. In one embodiment, the polypeptide sequences of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493 are encoded by the full length cDNA sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480,  
30    4482, 4484, 4486, 4488, 4490, 4492, and 4494, respectively.

Nucleic acids encoding polypeptides and proteins that are variants of the polypeptides and proteins encoded by the present nucleic acids and related cDNA and genes are also within the scope of the invention. The variants differ from wild-type protein in having one or more amino acid substitutions that either enhance, add, or diminish a biological activity of the wild-type protein. Once the amino acid change is selected, a nucleic acid encoding that variant is constructed according to the invention.

The following detailed description discloses how to obtain or make full-length cDNA and human genes corresponding to the nucleic acids, how to express these nucleic acids and genes, how to identify structural motifs of the genes, how to identify the function of a protein encoded by a gene corresponding to an nucleic acid, how to use nucleic acids as probes in mapping and in tissue profiling, how to use the corresponding polypeptides and proteins to raise antibodies, and how to use the nucleic acids, polypeptides, and proteins for diagnostic purposes.

The sequences disclosed herein have been found to be differentially expressed in colon cancer cell lines and/or colon cancer tissue, and thus are useful for determining the presence of colon cancer in a cell or tissue sample. The present sequences also have utility for determining the presence or state of other types of cancer.

Accordingly, a preferred aspect of the present invention relates to nucleic acids differentially expressed in tumor cells or tissue, especially colon cancer tissue or cells, polypeptides encoded by such nucleic acids, and antibodies immunoreactive with these polypeptides, and preparations of such compositions. Moreover, the present invention provides diagnostic and therapeutic assays and reagents for detecting and treating disorders involving, for example, expression of the subject nucleic acids.

#### I. General

This invention relates to compositions and methods for identifying and/or classifying cancerous cells present in a human tumors, particularly in solid tumors, e.g., carcinomas and sarcomas, such as, for example, breast or colon cancers. In its broadest aspect, the method uses nucleic acids that are differentially expressed in cancer cell lines and/or cancer tissue, compared with related normal cells or tissue, and using them to identify or classify tumor cells by the upregulation and/or downregulation of expression of particular genes, an event which is implicated in tumorigenesis.

Upregulation or increased expression of certain genes such as oncogenes, act to promote malignant growth. Downregulation or decreased expression of genes, such as tumor suppressor

genes, also promotes malignant growth. Thus, alteration in the expression of either type of gene is a potential diagnostic indicator for determining whether a subject is at risk of developing or has cancer, e.g., colon cancer.

Accordingly, in one aspect, the invention also provides biomarkers, such as nucleic acid  
5 markers, for human tumor cells and tissue, particularly for colon cancer cells and tissue. The invention also provides proteins encoded by these nucleic acid markers. The invention also features methods for identifying drugs useful for treatment of such cancer cells, and for treatment of a cancerous condition, such as colon cancer. Unlike prior methods, the invention provides a means for identifying cancer cells at an early stage of development, so that premalignant cells  
10 can be identified prior to their spreading throughout the human body. This allows early detection of potentially cancerous conditions, and treatment of those cancerous conditions prior to spread of the cancerous cells throughout the body, or prior to development of an irreversible cancerous condition.

## II. Definitions

15 For convenience, the meaning of certain terms and phrases used in the specification, examples, and appended claims, are provided below.

The term “an aberrant expression”, as applied to a nucleic acid of the present invention, refers to level of expression of that nucleic acid which differs from the level of expression of that nucleic acid in healthy tissue, or which differs from the activity of the polypeptide present in a  
20 healthy subject. An activity of a polypeptide can be aberrant because it is stronger than the activity of its native counterpart. Alternatively, an activity can be aberrant because it is weaker or absent relative to the activity of its native counterpart. An aberrant activity can also be a change in the activity; for example, an aberrant polypeptide can interact with a different target peptide. A cell can have an aberrant expression level of a gene due to overexpression or  
25 underexpression of that gene.

The term “agonist”, as used herein, is meant to refer to an agent that mimics or upregulates (e.g., potentiates or supplements) the bioactivity of a protein. An agonist can be a wild-type protein or derivative thereof having at least one bioactivity of the wild-type protein. An agonist can also be a compound that upregulates expression of a gene or which increases at  
30 least one bioactivity of a protein. An agonist can also be a compound which increases the interaction of a polypeptide with another molecule, e.g., a target peptide or nucleic acid.



The term “allele”, which is used interchangeably herein with “allelic variant”, refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for that gene or allele. When a subject has two different alleles of a gene, the subject is said to be heterozygous for the gene. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and/or insertions of nucleotides. An allele of a gene can also be a form of a gene containing mutations.

The term “allelic variant of a polymorphic region of a gene” refers to a region of a gene having one of several nucleotide sequences found in that region of the gene in other individuals.

The term “antagonist” as used herein is meant to refer to an agent that downregulates (e.g., suppresses or inhibits) at least one bioactivity of a protein. An antagonist can be a compound which inhibits or decreases the interaction between a protein and another molecule, e.g., a target peptide or enzyme substrate. An antagonist can also be a compound that downregulates expression of a gene or which reduces the amount of expressed protein present.

The term “antibody” as used herein is intended to include whole antibodies, e.g., of any isotype (IgG, IgA, IgM, IgE, etc), and includes fragments thereof which are also specifically reactive with a vertebrate, e.g., mammalian, protein. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. Thus, the term includes segments of proteolytically-cleaved or recombinantly-prepared portions of an antibody molecule that are capable of selectively reacting with a certain protein. Nonlimiting examples of such proteolytic and/or recombinant fragments include Fab, F(ab')<sub>2</sub>, Fab', Fv, and single chain antibodies (scFv) containing a V[L] and/or V[H] domain joined by a peptide linker. The scFv's may be covalently or non-covalently linked to form antibodies having two or more binding sites. The subject invention includes polyclonal, monoclonal, or other purified preparations of antibodies and recombinant antibodies.

The phenomenon of “apoptosis” is well known, and can be described as a programmed death of cells. As is known, apoptosis is contrasted with “necrosis”, a phenomenon when cells die as a result of being killed by a toxic material, or other external effect. Apoptosis involves chromatic condensation, membrane blebbing, and fragmentation of DNA, all of which are generally visible upon microscopic examination.

A disease, disorder, or condition “associated with” or “characterized by” an aberrant expression of a nucleic acid refers to a disease, disorder, or condition in a subject which can be statistically correlated with the expression of a nucleic acid.

As used herein the term “bioactive fragment of a polypeptide” refers to a fragment of a full-length polypeptide, wherein the fragment specifically agonizes (mimics) or antagonizes (inhibits) the activity of a wild-type polypeptide. The bioactive fragment preferably is a fragment capable of interacting with at least one other molecule, e.g., protein, small molecule, or DNA, which a full length protein can bind.

“Biological activity” or “bioactivity” or “activity” or “biological function”, which are used interchangeably, herein mean an effector or antigenic function that is directly or indirectly performed by a polypeptide (whether in its native or denatured conformation), or by any subsequence thereof. Biological activities include binding to polypeptides, binding to other proteins or molecules, activity as a DNA binding protein, as a transcription regulator, ability to bind damaged DNA, etc. A bioactivity can be modulated by directly affecting the subject polypeptide. Alternatively, a bioactivity can be altered by modulating the level of the polypeptide, such as by modulating expression of the corresponding gene.

The term “biomarker” refers a biological molecule, e.g., a nucleic acid, including DNA, cDNA, RNA, mRNA, tRNA, or rRNA, peptide, polypeptide, protein, hormone, etc., whose presence or concentration can be detected and correlated with a known condition, such as a disease state.

“Cells,” “host cells”, or “recombinant host cells” are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A “chimeric polypeptide” or “fusion polypeptide” is a fusion of a first amino acid sequence encoding one of the subject polypeptides with a second amino acid sequence defining a domain (e.g., polypeptide portion) foreign to and not substantially homologous with any domain of the subject polypeptide. A chimeric polypeptide may present a foreign domain which is found (albeit in a different polypeptide) in an organism which also expresses the first polypeptide, or it may be an “interspecies,” “intergenic,” etc., fusion of polypeptide structures expressed by different kinds of organisms. In general, a fusion polypeptide can be represented by the general

formula  $(X)_n-(Y)_m-(Z)_n$ , wherein Y represents a portion of the subject polypeptide, and X and Z are each independently absent or represent amino acid sequences which are not related to the native sequence found in an organism, or which are not found as a polypeptide chain contiguous with the subject sequence, where m is an integer greater than or equal to one, and each  
5 occurrence of n is, independently, 0 or an integer greater than or equal to 1 (n and m are preferably no greater than 5 or 10).

A “delivery complex” shall mean a targeting means (e.g., a molecule that results in higher affinity binding of a nucleic acid, protein, polypeptide or peptide to a target cell surface and/or increased cellular or nuclear uptake by a target cell). Examples of targeting means  
10 include: sterols (e.g., cholesterol), lipids (e.g., a cationic lipid, virosome or liposome), viruses (e.g., adenovirus, adeno-associated virus, and retrovirus), or target cell-specific binding agents (e.g., ligands recognized by target cell specific receptors). Preferred complexes are sufficiently stable *in vivo* to prevent significant uncoupling prior to internalization by the target cell. However, the complex is cleavable under appropriate conditions within the cell so that the  
15 nucleic acid, protein, polypeptide or peptide is released in a functional form.

As is well known, genes or a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term “DNA sequence encoding a  
20 polypeptide” may thus refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences may exist between individual organisms, which are called alleles. Such allelic differences may or may not result in differences in amino acid sequence of the encoded polypeptide yet still encode a polypeptide with the same biological activity.

25 The term “equivalent” is understood to include nucleotide sequences encoding functionally equivalent polypeptides. Equivalent nucleotide sequences will include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include sequences that differ from the nucleotide sequence of the nucleic acids shown in SEQ ID NOs: 1-4494 due to the degeneracy of the genetic code.

30 As used herein, the terms “gene”, “recombinant gene”, and “gene construct” refer to a nucleic acid of the present invention associated with an open reading frame, including both exon and, optionally, intron sequences.

A “recombinant gene” refers to nucleic acid encoding a polypeptide and comprising exon sequences, though it may optionally include intron sequences which are derived from, for example, a related or unrelated chromosomal gene. The term “intron” refers to a DNA sequence present in a given gene which is not translated into protein and is generally found between exons.

5           The term “growth” or “growth state” of a cell refers to the proliferative state of a cell as well as to its differentiative state. Accordingly, the term refers to the phase of the cell cycle in which the cell is, e.g., G<sub>0</sub>, G<sub>1</sub>, G<sub>2</sub>, or prophase, metaphase, or telophase, or anaphase, as well as to its state of differentiation, e.g., undifferentiated, partially differentiated, or fully differentiated. Without wanting to be limited, differentiation of a cell is usually accompanied by a decrease in  
10   the proliferative rate of a cell.

          “Homology” or “identity” or “similarity” refers to sequence similarity between two peptides or between two nucleic acid molecules, with identity being a more strict comparison. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is  
15   occupied by the same base or amino acid, then the molecules are identical at that position. A degree of homology or similarity or identity between nucleic acid sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. A degree of identity of amino acid sequences is a function of the number of identical amino acids at positions shared by the amino acid sequences. A degree of homology or similarity of amino acid  
20   sequences is a function of the number of amino acids, i.e., structurally related, at positions shared by the amino acid sequences. An “unrelated” or “non-homologous” sequence shares less than 40% identity, though preferably less than 25% identity, with one of the sequences of the present invention.

          The term “percent identical” refers to sequence identity between two amino acid  
25   sequences or between two nucleotide sequences. Identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules  
30   can be referred to as homologous (similar) at that position. Expression as a percentage of homology, similarity, or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. Various alignment algorithms and/or programs may be used, including FASTA, BLAST, or ENTREZ. FASTA and BLAST are

available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with, e.g., default settings. ENTREZ is available through the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Md. In one embodiment, the percent identity of two sequences can be  
5 determined by the GCG program with a gap weight of 1, e.g., each amino acid gap is weighted as if it were a single amino acid or nucleotide mismatch between the two sequences.

Other techniques for alignment are described in Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USA. Preferably, an  
10 alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See Meth. Mol. 70-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAR computer. MPSRCH uses a Smith-Waterman algorithm to  
15 score sequences on a massively parallel computer. This approach improves ability to pick up distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Nucleic acid-encoded amino acid sequences can be used to search both protein and DNA databases.

Databases with individual sequences are described in Methods in Enzymology, ed.  
20 Doolittle, *supra*. Databases include, for example, Genbank, EMBL, and DNA Database of Japan (DDBJ).

Preferred nucleic acids have a sequence at least 70%, and more preferably 80% identical and more preferably 90% and even more preferably at least 95% identical to a nucleic acid sequence of a sequence shown in one of SEQ ID NOS: 1-4494. Nucleic acids at least 90%, more  
25 preferably 95%, and most preferably at least about 98-99% identical with a nucleic sequence represented in one of SEQ ID NOS: 1-4494 are of course also within the scope of the invention. In preferred embodiments, the nucleic acid is mammalian.

The term "interact" as used herein is meant to include detectable interactions (e.g., biochemical interactions) between molecules, such as interaction between protein-protein,  
30 protein-nucleic acid, nucleic acid-nucleic acid, and protein-small molecule or nucleic acid-small molecule in nature. Examples of interactions between protein-protein, protein-nucleic acid, nucleic acid-nucleic acid, and protein-small molecule or nucleic acid-small molecule can include binding, modifying, cleaving, processing, or catalyzing.

The term “isolated” as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an “isolated nucleic acid” is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term “isolated” is also used herein to refer to polypeptides which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides.

The terms “modulated” and “differentially regulated” as used herein refer to both upregulation (i.e., activation or stimulation e.g., by agonizing or potentiating) and downregulation (i.e., inhibition or suppression e.g., by antagonizing, decreasing or inhibiting).

The term “mutated gene” refers to an allelic form of a gene, which is capable of altering the phenotype of a subject having the mutated gene relative to a subject which does not have the mutated gene. If a subject must be homozygous for this mutation to have an altered phenotype, the mutation is said to be recessive. If one copy of the mutated gene is sufficient to alter the genotype of the subject, the mutation is said to be dominant. If a subject has one copy of the mutated gene and has a phenotype that is intermediate between that of a homozygous and that of a heterozygous subject (for that gene), the mutation is said to be co-dominant.

The designation “N”, where it appears in the accompanying Sequence Listing, indicates that the identity of the corresponding nucleotide is unknown. “N” should therefore not necessarily be interpreted as permitting substitution with any nucleotide, e.g., A, T, C, or G, but rather as holding the place of a nucleotide whose identity has not been conclusively determined.

The “non-human animals” of the invention include mammals such as rodents, non-human primates, sheep, dog, cow, pigs, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation. The term “chimeric animal” is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant gene is expressed in some but not all cells of the animal. The term “tissue-specific chimeric

animal” indicates that one of the recombinant genes is present and/or expressed or disrupted in some tissues but not others.

As used herein, the term “nucleic acid” refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides. ESTs, chromosomes, cDNAs, mRNAs, and rRNAs are representative examples of molecules that may be referred to as nucleic acids.

The term “nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO. x” refers to the nucleotide sequence of the complementary strand of a nucleic acid strand having SEQ ID NO. x. The term “complementary strand” is used herein interchangeably with the term “complement”. The complement of a nucleic acid strand can be the complement of a coding strand or the complement of a non-coding strand. As used herein, a “complementary strand” to SEQ ID NO. x is a nucleic acid sequence which hybridizes under stringent conditions to SEQ ID NO. x.

The term “polymorphism” refers to the coexistence of more than one form of a gene or portion (e.g., allelic variant) thereof. A portion of a gene of which there are at least two different forms, i.e., two different nucleotide sequences, is referred to as a “polymorphic region of a gene”. A polymorphic region can be a single nucleotide, the identity of which differs in different alleles. A polymorphic region can also be several nucleotides long.

A “polymorphic gene” refers to a gene having at least one polymorphic region.

As used herein, the term “promoter” means a DNA sequence that regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in cells. The term encompasses “tissue specific” promoters, i.e., promoters which effect expression of the selected DNA sequence only in specific cells (e.g., cells of a specific tissue). The term also covers so-called “leaky” promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well. The term also encompasses non-tissue specific promoters and promoters that constitutively expressed or that are inducible (i.e., expression levels can be controlled).

The terms “protein”, “polypeptide”, and “peptide” are used interchangeably herein when referring to a gene product.

The term “recombinant protein” refers to a polypeptide of the present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding a polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase “derived from”, with respect to a recombinant gene, is meant to include within the meaning of “recombinant protein” those proteins having an amino acid sequence of a native polypeptide, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions (including truncation) of a naturally occurring form of the polypeptide.

“Small molecule” as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic (carbon-containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention to identify compounds that modulate a bioactivity.

As used herein, the term “specifically hybridizes” or “specifically detects” refers to the ability of a nucleic acid molecule of the invention to hybridize to at least a portion of, for example approximately 6, 12, 15, 20, 30, 50, 100, 150, 200, 300, 350, 400, 500, 750, or 1000 contiguous nucleotides of a nucleic acid designated in any one of SEQ ID Nos: 1-4494, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5% background hybridization to a cellular nucleic acid (e.g., mRNA or genomic DNA) encoding a different protein. In preferred embodiments, the oligonucleotide probe detects only a specific nucleic acid, e.g., it does not substantially hybridize to similar or related nucleic acids, or complements thereof.

“Transcriptional regulatory sequence” is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the genes is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally occurring forms of the polypeptide.



As used herein, the term “transfection” means the introduction of a nucleic acid, e.g., via an expression vector, into a recipient cell by nucleic acid-mediated gene transfer.

“Transformation”, as used herein, refers to a process in which a cell’s genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of a polypeptide or, in the case of anti-sense expression from the transferred gene, the expression of the target gene is disrupted.

The term “treating” as used herein is intended to encompass curing as well as ameliorating at least one symptom of the condition or disease.

The term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as “expression vectors”. In general, expression vectors of utility in recombinant DNA techniques are often in the form of “plasmids” which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, “plasmid” and “vector” are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

The term “wild-type allele” refers to an allele of a gene which, when present in two copies in a subject results in a wild-type phenotype. There can be several different wild-type alleles of a specific gene, since certain nucleotide changes in a gene may not affect the phenotype of a subject having two copies of the gene with the nucleotide changes.

### III. Nucleic Acids of the Present Invention

As described below, one aspect of the invention pertains to isolated nucleic acids, variants, and/or equivalents of such nucleic acids.

Nucleic acids of the present invention have been identified as differentially expressed in tumor cells, e.g., colon cancer-derived cell lines and colon cancer tissue (relative to the expression levels in normal cells or tissue, e.g., normal colon tissue and/or normal non-colon tissue). The present differentially expressed sequences comprise SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos.

1-1103, even more preferably SEQ ID Nos. 1-503, or sequence complementary thereto. In another embodiment, the invention comprises sequences which hybridize under stringent conditions with any of the sequences of SEQ ID Nos 1-4494. In a preferred aspect, sequences of the invention hybridize to SEQ ID Nos 1-4494 with about 50% identity, preferably about 70% identity, more preferably about 90% identity, and still more preferably about 100% identity. In preferred embodiments, the subject nucleic acids are differentially expressed by at least a factor of two, preferably at least a factor of five, even more preferably at least a factor of twenty, still more preferably at least a factor of fifty. Preferred nucleic acids are those sequences identified as differentially expressed both in colon cancer tissue and colon cancer cell lines. In preferred embodiments, nucleic acids of the present invention are upregulated in tumor cells, especially colon cancer tissue and/or colon cancer-derived cell lines. In another embodiment, nucleic acids of the present invention are downregulated in tumor cells, especially colon cancer tissue and/or colon cancer-derived cell lines.

Genes which are upregulated, such as oncogenes, or downregulated, such as tumor suppressors, in aberrantly proliferating cells can be used as targets for diagnostic or therapeutic applications. For example, upregulation of the *cdc2* gene induces mitosis. Overexpression of the *myt1* gene, a mitotic deactivator, negatively regulates the activity of *cdc2*. Aberrant proliferation may thus be induced either by upregulating *cdc2* or by downregulating *myt1*. Similarly, downregulation of tumor suppressors such as p53 and Rb have been implicated in tumorigenesis.

Particularly preferred polypeptides are those that are encoded by nucleic acid sequences at least about 70%, 75%, 80%, 90%, 95%, 97%, or 98% similar to a nucleic acid sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. Preferably, the nucleic acid includes all or a portion (e.g., at least about 10, at least about 15, at least about 25, or at least about 40 nucleotides) of the nucleotide sequence corresponding to the nucleic acid of SEQ ID Nos. 1-1103, most preferably SEQ ID Nos. 1-503, or a sequence complementary thereto.

Still other preferred nucleic acids of the present invention encode a polypeptide comprising at least a portion of a polypeptide encoded by one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. For example, preferred nucleic acid molecules for use as probes/primers or antisense molecules (i.e., noncoding nucleic acid molecules) can comprise at least about 10, 20, 30, 50, 60, 70, 80, 90, or 100 base pairs in length up to the length of the complete sequence of any of SEQ ID Nos 1-4494. Coding nucleic

acid molecules can comprise, for example, from about 50, 60, 70, 80, 90, or 100 base pairs up to the full length of the entire sequence of any of SEQ ID Nos 1-4494.

Another aspect of the invention provides a nucleic acid which hybridizes under low, medium, or high stringency conditions to a nucleic acid sequence represented by one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto. Appropriate stringency conditions which promote DNA hybridization, for example, about 6.0 x sodium chloride/sodium citrate (SSC) at about 45 °C, followed by a wash of about 2.0 x SSC at about 50°C, are known to those skilled in the art or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-12.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at about 50°C to a high stringency of about 0.2 x SSC at about 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22 °C, to high stringency conditions at about 65 °C. Both temperature and salt may be varied, or temperature or salt concentration may be held constant while the other variable is changed. In a preferred embodiment, a nucleic acid of the present invention will bind to one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, under moderately stringent conditions, for example at about 2.0 x SSC and about 40°C. In a particularly preferred embodiment, a nucleic acid of the present invention will bind to one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, under high stringency conditions.

In one embodiment, the invention provides nucleic acids which hybridize under low stringency conditions of about 6 x SSC at about room temperature followed by a wash at about 2 x SSC at about room temperature.

In another embodiment, the invention provides nucleic acids which hybridize under high stringency conditions of about 2 x SSC at about 65 °C followed by a wash at about 0.2 x SSC at about 65 °C.

Nucleic acids having a sequence that differs from the nucleotide sequences shown in one of SEQ ID Nos. 1-1103, preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, due to degeneracy in the genetic code, are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides (i.e., a peptide having equivalent or similar biological activity) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and

CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence of a polypeptide. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject polypeptides will exist among mammals. One skilled in the art will appreciate that these variations in one or more nucleotides (e.g., up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides having an activity of a polypeptide may exist among individuals of a given species due to natural allelic variation.

Also within the scope of the invention are nucleic acids encoding splicing variants of proteins encoded by a nucleic acid of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, or natural homologs of such proteins. Such homologs can be cloned by hybridization or PCR, as further described herein.

The polynucleotide sequence may also encode for a leader sequence, e.g., the natural leader sequence or a heterologous leader sequence, for a subject polypeptide. For example, the desired DNA sequence may be fused in the same reading frame to a DNA sequence which aids in expression and secretion of the polypeptide from the host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of the polypeptide from the cell. The protein having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form the mature form of the protein.

The polynucleotide of the present invention may also be fused in frame to a marker sequence, also referred to herein as "Tag sequence" encoding a "Tag peptide", which allows for marking and/or purification of the present invention. In a preferred embodiment, the marker sequence is a hexahistidine tag, e.g., supplied by a PQE-9 vector. Numerous other Tag peptides are available commercially. Other frequently used Tags include myc-epitopes (e.g., see Ellison et al. (1991) J Biol Chem 266:21150-21157) which includes a 10-residue sequence from c-myc, the pFLAG system (International Biotechnologies, Inc.), the pEZZ-protein A system (Pharmacia, NJ), and a 16 amino acid portion of the Haemophilus influenza hemagglutinin protein. Furthermore, any polypeptide can be used as a Tag so long as a reagent, e.g., an antibody interacting specifically with the Tag polypeptide is available or can be prepared or identified.

As indicated by the examples set out below, nucleic acids can be obtained from mRNA present in any of a number of eukaryotic cells or tissue, e.g., and are preferably obtained from metazoan cells or tissue, more preferably from vertebrate cells or tissue, and even more preferably from mammalian cells and tissue, and most preferably from human cells or tissue. It

also is possible to obtain nucleic acids of the present invention from genomic DNA from both adults and embryos. For example, a gene can be cloned from either a cDNA or a genomic library in accordance with protocols generally known to persons skilled in the art. cDNA can be obtained by isolating total mRNA from a cell, e.g., a vertebrate cell, a mammalian cell, or a human cell, including embryonic cells. Double stranded cDNAs can then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. The gene can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention.

The invention includes within its scope a polynucleotide having the nucleotide sequence of nucleic acid obtained from this biological material, wherein the nucleic acid hybridizes under stringent conditions (at least about 4 x SSC at 65 °C, or at least about 4 x SSC at 42 °C; see, for example, U.S. Patent No. 5,707,829, incorporated herein by reference) with at least 15 contiguous nucleotides of at least one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. By this is intended that when at least 15 contiguous nucleotides of one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 is used as a probe, the probe will preferentially hybridize with a gene or mRNA (of the biological material) comprising the complementary sequence, allowing the identification and retrieval of the nucleic acids of the biological material that uniquely hybridize to the selected probe. Probes from more than one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 will hybridize with the same gene or mRNA if the cDNA from which they were derived corresponds to one mRNA. Probes of more than 15 nucleotides can be used, but 15 nucleotides represents enough sequence for unique identification.

Because the present nucleic acids are cDNAs which represent partial mRNA transcripts, two or more nucleic acids of the invention may represent different regions of the same mRNA transcript and the same gene. Thus, if two or more of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 are identified as belonging to the same clone, then either sequence can be used to obtain the full-length mRNA or gene. Nucleic acid-related polynucleotides can also be isolated from cDNA libraries. These libraries are preferably prepared from mRNA of human colon cells, more preferably, human colon cancer specific tissue, designated as the 100-101, and 103-112 clones in Table 1. In another embodiment the nucleic acids are isolated from libraries prepared from normal colon specific tissue, designated herein as the 102 clones in Table 1. Alignment of SEQ ID Nos. 1-4470, 4472,

4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, as described above, indicated that a cell line or tissue source of a related protein or polynucleotide can also be used as a source of the nucleic acid-related cDNA.

Techniques for producing and probing nucleic acid sequence libraries are described, for example, in Sambrook et al., "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1989). The cDNA can be prepared by using primers based on a sequence from SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. In one embodiment, the cDNA library can be made from only poly-adenylated mRNA. Thus, poly-T primers can be used to prepare cDNA from the mRNA.

Alignment of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 can result in identification of a related polypeptide or polynucleotide. Some of the polynucleotides disclosed herein contains repetitive regions that were subject to masking during the search procedures. The information about the repetitive regions is discussed below.

Constructs of polynucleotides having sequences of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 can be generated synthetically. Alternatively, single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides is described by Stemmer et al, Gene (Amsterdam) (1995) 164(i):49-53. In this method, assembly PCR (the synthesis of long DNA sequences from large numbers of oligodeoxyribonucleotides (oligos)) is described. The method is derived from DNA shuffling (Stemmer, Nature (1994) 370:389-391), and does not rely on DNA ligase, but instead relies on DNA polymerase to build increasingly longer DNA fragments during the assembly process. For example, a 1.1-kb fragment containing the TEM-1 beta-lactamase-encoding gene (bla) can be assembled in a single reaction from a total of 56 oligos, each 40 nucleotides (nt) in length. The synthetic gene can be PCR amplified and cloned in a vector containing the tetracycline-resistance gene (Tc-R) as the sole selectable marker. Without relying on ampicillin (Ap) selection, 76% of the Tc-R colonies were Ap-R, making this approach a general method for the rapid and cost-effective synthesis of any gene.

#### IV. Identification of Functional and Structural Motifs of Novel Genes Using Art-Recognized Methods

Translations of the nucleotide sequence of the nucleic acids, cDNAs, or full genes can be aligned with individual known sequences. Similarity with individual sequences can be used to determine the activity of the polypeptides encoded by the polynucleotides of the invention. For

example, sequences that show similarity with a chemokine sequence may exhibit chemokine activities. Also, sequences exhibiting similarity with more than one individual sequence may exhibit activities that are characteristic of either or both individual sequences.

5 The full length sequences and fragments of the polynucleotide sequences of the nearest neighbors can be used as probes and primers to identify and isolate the full length sequence of the nucleic acid. The nearest neighbors can indicate a tissue or cell type to be used to construct a library for the full-length sequences of the nucleic acid.

10 Typically, the nucleic acids are translated in all six frames to determine the best alignment with the individual sequences. The sequences disclosed herein in the Sequence Listing are in a 5' to 3' orientation and translation in three frames can be sufficient (with a few specific exceptions as described in the Examples). These amino acid sequences are referred to, generally, as query sequences, which will be aligned with the individual sequences.

15 Nucleic acid sequences can be compared with known genes by any of the methods disclosed above. Results of individual and query sequence alignments can be divided into three categories: high similarity, weak similarity, and no similarity. Individual alignment results ranging from high similarity to weak similarity provide a basis for determining polypeptide activity and/or structure.

Parameters for categorizing individual results include: percentage of the alignment region length where the strongest alignment is found, percent sequence identity, and p value.

20 The percentage of the alignment region length is calculated by counting the number of residues of the individual sequence found in the region of strongest alignment. This number is divided by the total residue length of the query sequence to find a percentage.

25 Percent sequence identity is calculated by counting the number of amino acid matches between the query and individual sequence and dividing total number of matches by the number of residues of the individual sequence found in the region of strongest alignment. For the example above, the percent identity would be 10 matches divided by 11 amino acids, or approximately 90.9%.

30 P value is the probability that the alignment was produced by chance. For a single alignment, the p value can be calculated according to Karlin et al., Proc. Natl. Acad. Sci. 87: 2264 (1990) and Karlin et al., Proc. Natl. Acad. Sci. 90: (1993). The p value of multiple alignments using the same query sequence can be calculated using an heuristic approach

described in Altschul et al., Genet. 6:119(1994). Alignment programs such as BLAST program can calculate the p value.

5 The boundaries of the region where the sequences align can be determined according to Doolittle, Methods in Enzymology, *supra*; BLAST or FASTA programs; or by determining the area where the sequence identity is highest.

Another factor to consider for determining identity or similarity is the location of the similarity or identity. Strong local alignment can indicate similarity even if the length of alignment is short. Sequence identity scattered throughout the length of the query sequence also can indicate a similarity between the query and profile sequences.

#### 10 High Similarity

For the alignment results to be considered high similarity, the percent of the alignment region length, typically, is at least about 55% of total length query sequence; more typically, at least about 58%; even more typically; at least about 60% of the total residue length of the query sequence. Usually, percent length of the alignment region can be as much as about 62%; more  
15 usually, as much as about 64%; even more usually, as much as about 66%.

Further, for high similarity, the region of alignment, typically, exhibits at least about 75% of sequence identity; more typically, at least about 78%; even more typically; at least about 80% sequence identity. Usually, percent sequence identity can be as much as about 82%; more usually, as much as about 84%; even more usually, as much as about 86%.

20 The p value is used in conjunction with these methods. If high similarity is found, the query sequence is considered to have high similarity with a profile sequence when the p value is less than or equal to about  $10^{-2}$ ; more usually; less than or equal to about  $10^{-3}$  even more usually; less than or equal to about  $10^{-4}$ . More typically, the p value is no more than about  $10^{-5}$  more typically; no more than or equal to about  $10^{-10}$ ; even more typically; no more than or equal to  
25 about  $10^{-15}$  for the query sequence to be considered high similarity.

#### Weak Similarity

For the alignment results to be considered weak there is no minimum percent length of the alignment region no minimum length of alignment. A better showing of weak similarity is considered when the region of alignment is, typically, at least about 15 amino acid residues in  
30 length; more typically, at least about 20; even more typically; at least about 25 amino acid



residues in length. Usually, length of the alignment region can be as much as about 30 amino acid residues; more usually, as much as about 40; even more usually, as much as about 60 amino acid residues.

Further, for weak similarity, the region of alignment, typically, exhibits at least about  
5 35% of sequence identity; more typically, at least about 40%; even more typically; at least about 45% sequence identity. Usually, percent sequence identity can be as much as about 50%; more usually, as much as about 55%; even more usually, as much as about 60%.

If low similarity is found, the query sequence is considered to have weak similarity with a profile sequence when the p value is usually less than or equal to about  $10^{-2}$ ; more usually; less  
10 than or equal to about  $10^{-3}$  even more usually; less than or equal to about  $10^{-4}$ . More typically, the p value is no more than about  $10^{-5}$  more usually; no more than or equal to about  $10^{-10}$ ; even more usually; no more than or equal to about  $10^{-15}$  for the query sequence to be considered weak similarity.

#### Similarity Determined by Sequence Identity

Sequence identity alone can be used to determine similarity of a query sequence to an individual sequence and can indicate the activity of the sequence. Such an alignment, preferably, permits gaps to align sequences. Typically, the query sequence is related to the profile sequence if the sequence identity over the entire query sequence is at least about 15%; more typically, at least about 20%; even more typically, at least about 25%; even more typically, at least about  
20 50%. Sequence identity alone as a measure of similarity is most useful when the query sequence is usually, at least 80 residues in length; more usually, 90 residues; even more usually, at least 95 amino acid residues in length. More typically, similarity can be concluded based on sequence identity alone when the query sequence is preferably 100 residues in length; more preferably, 120 residues in length; even more preferably, 150 amino acid residues in length.

#### Determining Activity from Alignments with Profile and Multiple Aligned Sequences

Translations of the nucleic acids can be aligned with amino acid profiles that define either protein families or common motifs. Also, translations of the nucleic acids can be aligned to multiple sequence alignments (MSA) comprising the polypeptide sequences of members of protein families or motifs. Similarity or identity with profile sequences or MSAs can be used to  
30 determine the activity of the polypeptides encoded by nucleic acids or corresponding cDNA or genes. For example, sequences that show an identity or similarity with a chemokine profile or MSA can exhibit chemokine activities.

Profiles can be designed manually by (1) creating a MSA, which is an alignment of the amino acid sequence of members that belong to the family and (2) constructing a statistical representation of the alignment. Such methods are described, for example, in Birney et al., Nucl. Acid Res. 25(14): 2730-2739 (1996).

5           MSAs of some protein families and motifs are publicly available. For example, these include MSAs of 547 different families and motifs. These MSAs are described also in Sonnhammer et al., Proteins 28: 405-420 (1997). Other sources are also available in the world wide web. A brief description of these MSAs is reported in Pascarella et al., Prot. Eng. 9(3): 249-251 (1996).

10           Techniques for building profiles from MSAs are described in Sonnhammer et al., *supra*; Birney et al., *supra*; and Methods in Enzymology, vol. 266: "Computer Methods for Macromolecular Sequence Analysis," 1996, ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USA.

15           Similarity between a query sequence and a protein family or motif can be determined by (a) comparing the query sequence against the profile and/or (b) aligning the query sequence with the members of the family or motif.

20           Typically, a program such as Searchwise can be used to compare the query sequence to the statistical representation of the multiple alignment, also known as a profile. The program is described in Birney et al., *supra*. Other techniques to compare the sequence and profile are described in Sonnhammer et al., *supra* and Doolittle, *supra*.

          Next, methods described by Feng et al., J. Mol. Evol. 25:351-360 (1987) and Higgins et al., CABIOS 5:151-153 (1989) can be used to align the query sequence with the members of a family or motif, also known as a MSA. Computer programs, such as PILEUP, can be used. See Feng et al., *infra*.

25           The following factors are used to determine if a similarity between a query sequence and a profile or MSA exists: (1) number of conserved residues found in the query sequence, (2) percentage of conserved residues found in the query sequence, (3) number of frameshifts, and (4) spacing between conserved residues.

30           Some alignment programs that both translate and align sequences can make any number of frameshifts when translating the nucleotide sequence to produce the best alignment. The fewer frameshifts needed to produce an alignment, the stronger the similarity or identity between

the query and profile or MSAs. For example, a weak similarity resulting from no frameshifts can be a better indication of activity or structure of a query sequence, than a strong similarity resulting from two frameshifts.

Preferably, three or fewer frameshifts are found in an alignment; more preferably two or fewer frameshifts; even more preferably, one or fewer frameshifts; even more preferably, no frameshifts are found in an alignment of query and profile or MSAs.

Conserved residues are those amino acids that are found at a particular position in all or some of the family or motif members. For example, most known chemokines contain four conserved cysteines. Alternatively, a position is considered conserved if only a certain class of amino acids is found in a particular position in all or some of the family members. For example, the N-terminal position may contain a positively charged amino acid, such as lysine, arginine, or histidine.

Typically, a residue of a polypeptide is conserved when a class of amino acids or a single amino acid is found at a particular position in at least about 40% of all class members; more typically, at least about 50%; even more typically, at least about 60% of the members. Usually, a residue is conserved when a class or single amino acid is found in at least about 70% of the members of a family or motif; more usually, at least about 80%; even more usually, at least about 90%; even more usually, at least about 95%.

A residue is considered conserved when three unrelated amino acids are found at a particular position in the some or all of the members; more usually, two unrelated amino acids. These residues are conserved when the unrelated amino acids are found at particular positions in at least about 40% of all class member, more typically, at least about 50%; even more typically, at least about 60% of the members. Usually, a residue is conserved when a class or single amino acid is found in at least about 70% of the members of a family or motif more usually, at least about 80%; even more usually, at least about 90%; even more usually, at least about 95%.

A query sequence has similarity to a profile or MSA when the query sequence comprises at least about 25% of the conserved residues of the profile or MSA; more usually, at least about 30%; even more usually; at least about 40%. Typically, the query sequence has a stronger similarity to a profile sequence or MSA when the query sequence comprises at least about 45% of the conserved residues of the profile or MSA more typically, at least about 50%; even more typically; at least about 55%.

V. Probes and Primers

The nucleotide sequences determined from the cloning of genes from tumor cells, especially colon cancer cell lines and tissues will further allow for the generation of probes and primers designed for identifying and/or cloning homologs in other cell types, e.g., from other tissues, as well as homologs from other mammalian organisms. Nucleotide sequences useful as probes/primers may include all or a portion of the sequences listed in SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or sequences complementary thereto or sequences which hybridize under stringent conditions to all or a portion of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. For instance, the present invention also provides a probe/primer comprising a substantially purified oligonucleotide, which oligonucleotide comprising a nucleotide sequence that hybridizes under stringent conditions to at least approximately 12, preferably 25, more preferably 40, 50, or 75 consecutive nucleotides up to the full length of the sense or anti-sense sequence selected from the group consisting of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, or naturally occurring mutants thereof. For instance, primers based on a nucleic acid represented in SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and even still more preferred SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, can be used in PCR reactions to clone homologs of that sequence.

In yet another embodiment, the invention provides probes/primers comprising a nucleotide sequence that hybridizes under moderately stringent conditions to at least approximately 12, 16, 25, 40, 50 or 75 consecutive nucleotides up to the full length of the sense or antisense sequence selected from the group consisting of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or naturally occurring mutants thereof.

In particular, these probes are useful because they provide a method for detecting mutations in wild-type genes of the present invention. Nucleic acid probes which are complementary to a wild-type gene of the present invention and can form mismatches with mutant genes are provided, allowing for detection by enzymatic or chemical cleavage or by shifts in electrophoretic mobility. Likewise, probes based on the subject sequences can be used to

detect transcripts or genomic sequences encoding the same or homologous proteins, for use, for example, in prognostic or diagnostic assays. In preferred embodiments, the probe further comprises a label group attached thereto and able to be detected, e.g., the label group is selected from radioisotopes, fluorescent compounds, chemiluminescent compounds, enzymes, and enzyme co-factors.

Full-length cDNA molecules comprising the disclosed nucleic acids are obtained as follows. In a preferred embodiment, the invention provides the full length cDNA sequence of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. A subject nucleic acid or a portion thereof comprising at least about 12, 15, 18, or 20 nucleotides up to the full length of a sequence represented in SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, may be used as a hybridization probe to detect hybridizing members of a cDNA library using probe design methods, cloning methods, and clone selection techniques as described in U.S. Patent No. 5,654,173, "Secreted Proteins and Polynucleotides Encoding Them," incorporated herein by reference. Libraries of cDNA may be made from selected tissues, such as normal or tumor tissue, or from tissues of a mammal treated with, for example, a pharmaceutical agent. Preferably, the tissue is the same as that used to generate the nucleic acids, as both the nucleic acid and the cDNA represent expressed genes. Most preferably, the cDNA library is made from the biological material described herein in the Examples. Alternatively, many cDNA libraries are available commercially. (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed. (Cold Spring Harbor Press, Cold Spring Harbor, NY 1989). The choice of cell type for library construction may be made after the identity of the protein encoded by the nucleic acid-related gene is known. This will indicate which tissue and cell types are likely to express the related gene, thereby containing the mRNA for generating the cDNA.

Members of the library that are larger than the nucleic acid, and preferably that contain the whole sequence of the native message, may be obtained. To confirm that the entire cDNA has been obtained, RNA protection experiments may be performed as follows. Hybridization of a full-length cDNA to an mRNA may protect the RNA from RNase degradation. If the cDNA is not full length, then the portions of the mRNA that are not hybridized may be subject to RNase degradation. This may be assayed, as is known in the art, by changes in electrophoretic mobility on polyacrylamide gels, or by detection of released monoribonucleotides. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed. (Cold Spring Harbor Press, Cold Spring Harbor, NY 1989). In order to obtain additional sequences 5' to the end of a partial cDNA, 5'

RACE (PCR Protocols: A Guide to Methods and Applications (Academic Press, Inc. 1990)) may be performed.

Genomic DNA may be isolated using nucleic acids in a manner similar to the isolation of full-length cDNAs. Briefly, the nucleic acids, or portions thereof, may be used as probes to libraries of genomic DNA. Preferably, the library is obtained from the cell type that was used to generate the nucleic acids. Most preferably, the genomic DNA is obtained from the biological material described herein in the Example. Such libraries may be in vectors suitable for carrying large segments of a genome, such as P1 or YAC, as described in detail in Sambrook et al., 9.4-9.30. In addition, genomic sequences can be isolated from human BAC libraries, which are commercially available from Research Genetics, Inc., Huntsville, Alabama, USA, for example. In order to obtain additional 5' or 3' sequences, chromosome walking may be performed, as described in Sambrook et al., such that adjacent and overlapping fragments of genomic DNA are isolated. These may be mapped and pieced together, as is known in the art, using restriction digestion enzymes and DNA ligase.

Using the nucleic acids of the invention, corresponding full length genes can be isolated using both classical and PCR methods to construct and probe cDNA libraries. Using either method, Northern blots, preferably, may be performed on a number of cell types to determine which cell lines express the gene of interest at the highest rate.

Classical methods of constructing cDNA libraries in Sambrook et al., supra. With these methods, cDNA can be produced from mRNA and inserted into viral or expression vectors. Typically, libraries of mRNA comprising poly(A) tails can be produced with poly(T) primers. Similarly, cDNA libraries can be produced using the instant sequences as primers.

PCR methods may be used to amplify the members of a cDNA library that comprise the desired insert. In this case, the desired insert may contain sequence from the full length cDNA that corresponds to the instant nucleic acids. Such PCR methods include gene trapping and RACE methods.

Gene trapping may entail inserting a member of a cDNA library into a vector. The vector then may be denatured to produce single stranded molecules. Next, a substrate-bound probe, such a biotinylated oligo, may be used to trap cDNA inserts of interest. Biotinylated probes can be linked to an avidin-bound solid substrate. PCR methods can be used to amplify the trapped cDNA. To trap sequences corresponding to the full length genes, the labeled probe sequence may be based on the nucleic acids of the invention, e.g., SEQ ID Nos. 1-1103, preferably SEQ

ID Nos. 1-503, or a sequence complementary thereto. Random primers or primers specific to the library vector can be used to amplify the trapped cDNA. Such gene trapping techniques are described in Gruber et al., PCT WO 95/04745 and Gruber et al., U.S. Pat. No. 5,500,356. Kits are commercially available to perform gene trapping experiments from, for example, Life Technologies, Gaithersburg, Maryland, USA.

“Rapid amplification of cDNA ends,” or RACE, is a PCR method of amplifying cDNAs from a number of different RNAs. The cDNAs may be ligated to an oligonucleotide linker and amplified by PCR using two primers. One primer may be based on sequence from the instant nucleic acids, for which full length sequence is desired, and a second primer may comprise a sequence that hybridizes to the oligonucleotide linker to amplify the cDNA. A description of this method is reported, for example, in PCT Pub. No. WO 97/19110.

In preferred embodiments of RACE, a common primer may be designed to anneal to an arbitrary adaptor sequence ligated to cDNA ends (Apte and Siebert, Biotechniques, 15:890-893, 1993; Edwards et al., Nuc. Acids Res., 19:5227-5232, 1991). When a single gene-specific RACE primer is paired with the common primer, preferential amplification of sequences between the single gene specific primer and the common primer occurs. Commercial cDNA pools modified for use in RACE are available.

Another PCR-based method generates full-length cDNA library with anchored ends without specific knowledge of the cDNA sequence. The method uses lock-docking primers (I-VI), where one primer, poly TV (I-III) locks over the polyA tail of eukaryotic mRNA producing first strand synthesis and a second primer, polyGH (IV-VI) locks onto the polyC tail added by terminal deoxynucleotidyl transferase (TdT). This method is described, for example, in PCT Pub. No. WO 96/40998.

The promoter region of a gene generally is located 5' to the initiation site for RNA polymerase II. Hundreds of promoter regions contain the “TATA” box, a sequence such as TATTA or TATAA, which is sensitive to mutations. The promoter region can be obtained by performing 5' RACE using a primer from the coding region of the gene. Alternatively, the cDNA can be used as a probe for the genomic sequence, and the region 5' to the coding region is identified by “walking up.”

If the gene is highly expressed or differentially expressed, the promoter from the gene may be of use in a regulatory construct for a heterologous gene.

Once the full-length cDNA or gene is obtained, DNA encoding variants can be prepared by site-directed mutagenesis, described in detail in Sambrook 15.3-15.63. The choice of codon or nucleotide to be replaced can be based on the disclosure herein on optional changes in amino acids to achieve altered protein structure and/or function.

5 As an alternative method to obtaining DNA or RNA from a biological material, nucleic acid comprising nucleotides having the sequence of one or more nucleic acids of the invention can be synthesized. Thus, the invention encompasses nucleic acid molecules ranging in length from 12 nucleotides (corresponding to at least 12 contiguous nucleotides which hybridize under stringent conditions to or are at least 80% identical to a nucleic acid represented by one of SEQ  
10 ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto) up to a maximum length suitable for one or more biological manipulations, including replication and expression, of the nucleic acid molecule. The invention includes but is not limited to (a) nucleic acid having the size of a full gene, and comprising at  
15 least one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto; (b) the nucleic acid of (a) also comprising at least one additional gene, operably linked to permit expression of a fusion protein; (c) an expression vector comprising (a) or (b); (d) a plasmid comprising (a) or (b); and (e) a recombinant viral  
20 particle comprising (a) or (b). Construction of (c) can be accomplished as described below in part VI.

The sequence of a nucleic acid of the present invention is not limited and can be any sequence of A, T, G, and/or C (for DNA) and A, U, G, and/or C (for RNA) or modified bases thereof, including inosine and pseudouridine. The choice of sequence will depend on the desired  
25 function and can be dictated by coding regions desired, the intron-like regions desired, and the regulatory regions desired.

#### VI. Vectors Carrying Nucleic Acids of the Present Invention

The invention further provides plasmids and vectors, which can be used to express a gene in a host cell. The host cell may be any prokaryotic or eukaryotic cell. Thus, a nucleotide  
30 sequence derived from any one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, encoding all or a



selected portion of a protein, can be used to produce a recombinant form of an polypeptide via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures well known in the art.

Vectors that allow expression of a nucleic acid in a cell are referred to as expression vectors. Typically, expression vectors contain a nucleic acid operably linked to at least one transcriptional regulatory sequence. Regulatory sequences are art-recognized and are selected to direct expression of the subject nucleic acids. Transcriptional regulatory sequences are described in Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). In one embodiment, the expression vector includes a recombinant gene encoding a peptide having an agonistic activity of a subject polypeptide, or alternatively, encoding a peptide which is an antagonistic form of a subject polypeptide.

The choice of plasmid will depend on the type of cell in which propagation is desired and the purpose of propagation. Certain vectors are useful for amplifying and making large amounts of the desired DNA sequence. Other vectors are suitable for expression in cells in culture. Still other vectors are suitable for transfer and expression in cells in a whole animal or person. The choice of appropriate vector is well within the skill of the art. Many such vectors are available commercially. The nucleic acid or full-length gene is inserted into a vector typically by means of DNA ligase attachment to a cleaved restriction enzyme site in the vector. Alternatively, the desired nucleotide sequence may be inserted by homologous recombination in vivo. Typically this is accomplished by attaching regions of homology to the vector on the flanks of the desired nucleotide sequence. Regions of homology are added by ligation of oligonucleotides, or by polymerase chain reaction using primers comprising both the region of homology and a portion of the desired nucleotide sequence.

Nucleic acids or full-length genes are linked to regulatory sequences as appropriate to obtain the desired expression properties. These may include promoters (attached either at the 5' end of the sense strand or at the 3' end of the antisense strand), enhancers, terminators, operators, repressors, and inducers. The promoters may be regulated or constitutive. In some situations it may be desirable to use conditionally active promoters, such as tissue-specific or developmental stage-specific promoters. These are linked to the desired nucleotide sequence using the techniques described above for linkage to vectors. Any techniques known in the art may be used.

When any of the above host cells, or other appropriate host cells or organisms, are used to replicate and/or express the polynucleotides or nucleic acids of the invention, the resulting replicated nucleic acid, RNA, expressed protein or polypeptide, is within the scope of the invention as a product of the host cell or organism. The product is recovered by any appropriate means known in the art.

Once the gene corresponding to the nucleic acid is identified, its expression can be regulated in the cell to which the gene is native. For example, an endogenous gene of a cell can be regulated by an exogenous regulatory sequence as disclosed in U.S. Patent No. 5,641,670, "Protein Production and Protein Delivery."

A number of vectors exist for the expression of recombinant proteins in yeast (see, for example, Broach *et al* (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye, Academic Press, p. 83, incorporated by reference herein). In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, a polypeptide is produced recombinantly utilizing an expression vector generated by sub-cloning one of the nucleic acids represented in one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto.

The preferred mammalian expression vectors contain both prokaryotic sequences, to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The various methods employed in the preparation of plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989) Chapters 16 and 17.

When it is desirable to express only a portion of a gene, e.g., a truncation mutant, it may be necessary to add a start codon (ATG) to the oligonucleotide fragment containing the desired sequence to be expressed. It is well known in the art that a methionine at the N-terminal position can be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat *et al.*, (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller *et al.* (1987) *PNAS* 84:2718-1722). Therefore, removal of an N-terminal methionine, if desired, can be achieved either *in vivo* by expressing polypeptides in a host which produces

MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller *et al.*, *supra*).

Moreover, the nucleic acid constructs of the present invention can also be used as part of a gene therapy protocol to deliver nucleic acids such as antisense nucleic acids. Thus, another aspect of the invention features expression vectors for *in vivo* or *in vitro* transfection with an antisense oligonucleotide.

In addition to viral transfer methods, non-viral methods can also be employed to introduce a subject nucleic acid, e.g., a sequence represented by one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, into the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral targeting means of the present invention rely on endocytic pathways for the uptake of the subject nucleic acid by the targeted cell. Exemplary targeting means of this type include liposomal derived systems, polylysine conjugates, and artificial viral envelopes.

A nucleic acid of any of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, the corresponding cDNA, or the full-length gene may be used to express the partial or complete gene product. Appropriate nucleic acid constructs are purified using standard recombinant DNA techniques as described in, for example, Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. (Cold Spring Harbor Press, Cold Spring Harbor, New York), and under current regulations described in United States Dept. of HHS, National Institute of Health (NIH) Guidelines for Recombinant DNA research. The polypeptides encoded by the nucleic acid may be expressed in any expression system, including, for example, bacterial, yeast, insect, amphibian and mammalian systems. Suitable vectors and host cells are described, for example, in U.S. Patent No. 5,654,173.

Bacteria. Expression systems in bacteria include those described in Chang *et al.*, *Nature* (1978) 275:615, Goeddel *et al.*, *Nature* (1979) 281 :544, Goeddel *et al.*, *Nucleic Acids Rec.* (1980) 8:4057; EP 0 036,776, U.S. Patent No. 4,551,433, DeBoer *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1983) 80:2125, and Siebenlist *et al.*, *Cell* (1980) 20:269.

Yeast. Expression systems in yeast include those described in Hinnen *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1978) 75:1929; Ito *et al.*, *J. Bacteriol.* (1983) 153:163; Kurtz *et al.*, *Mol. Cell. Biol.* (1986) 6:142; Kunze *et al.*, *J. Basic Microbiol.* (1985) 25:141; Gleeson *et al.*, *J. Gen. Microbiol.* (1986) 132:3459, Roggenkamp *et al.*, *Mol. Gen. Genet.* (1986) 202:302) Das *et al.*, *J. Bacteriol.* (1984) 158:1165; De Louvencourt *et al.*, *J. Bacteriol.* (1983) 154:737, Van den Berg *et al.*, *Bio/Technology* (1990) 8:135; Kunze *et al.*, *J. Basic Microbiol.* (1985) 25:141; Cregg *et al.*, *Mol. Cell. Biol.* (1985) 5:3376, U.S. Patent Nos. 4,837,148 and 4,929,555; Beach and Nurse, *Nature* (1981) 300:706; Davidow *et al.*, *Curr. Genet.* (1985) 10:380, Gaillardin *et al.*, *Curr. Genet.* (1985) 10:49, Ballance *et al.*, *Biochem. Biophys. Res. Commun.* (1983) 112:284289; Tilburn *et al.*, *Gene* (1983) 26:205221, Yelton *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1984) 81:14701474, Kelly and Hynes, *EMBO J.* (1985) 4:475479; EP 0 244,234, and WO 91/00357.

Insect Cells. Expression of heterologous genes in insects is accomplished as described in U.S. Patent No. 4,745,051, Friesen *et al.*, (1986) "The Regulation of Baculovirus Gene Expression" in: *The Molecular Biology Of Baculoviruses* (W. Doerfler, ed.), EP 0 127,839, EP 0 155,476, and Vlak *et al.*, *J. Gen. Virol.* (1988) 69:765776, Miller *et al.*, *Ann. Rev. Microbiol.* (1988) 42:177, Carbonell *et al.*, *Gene* (1988) 73:409, Maeda *et al.*, *Nature* (1985) 315:592594, Lebacqz Verheyden *et al.*, *Mol. Cell. Biol.* (1988) 8:3129; Smith *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1985) 82:8404, Miyajima *et al.*, *Gene* (1987) 58:273; and Martinet *et al.*, *DNA* (1988) 7:99. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts are described in Luckow *et al.*, *Bio/Technology* (1988) 6:4755, Miller *et al.*, *Generic Engineering* (Setlow, J.K. *et al.* eds.), Vol. 8 (Plenum Publishing, 1986), pp. 277279, and Maeda *et al.*, *Nature*, (1985) 315:592-594.

Mammalian Cells. Mammalian expression is accomplished as described in Dijkema *et al.*, *EMBO J.* (1985) 4:761, Gorman *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1982) 79:6777, Boshart *et al.*, *Cell* (1985) 41:52 1 and U.S. Patent No. 4,399,216. Other features of mammalian expression are facilitated as described in Ham and Wallace, *Meth. Enz.* (1979) 58:44, Barnes and Sato, *Anal. Biochem.* (1980) 102:255, U.S. Patent Nos. 4,767,704, 4,657,866, 4,927,762, 4,560,655, WO 90/103430, WO 87/00195, and U.S. RE 30,985.

## VII. Therapeutic Nucleic Acid Constructs

One aspect of the invention relates to the use of the isolated nucleic acid, e.g., SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, in antisense therapy. As used herein, antisense therapy refers to

administration or *in situ* generation of oligonucleotide molecules or their derivatives which specifically hybridize (e.g., bind) under cellular conditions with the cellular mRNA and/or genomic DNA, thereby inhibiting transcription and/or translation of that gene. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA  
5 duplexes, through specific interactions in the major groove of the double helix. In general, antisense therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary  
10 to at least a unique portion of the cellular mRNA. Alternatively, the antisense construct is an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell, causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a subject nucleic acid. Such oligonucleotide probes are preferably modified oligonucleotides which are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and are  
15 therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphorothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol *et al.* (1988) *BioTechniques* 6:958-976; and Stein *et al.* (1988) *Cancer Res* 48:2659-2668. With  
20 respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the nucleotide sequence of interest, are preferred.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA. The antisense oligonucleotides will bind to the mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. In the  
25 case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable  
30 degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the mRNA, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most

efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at inhibiting translation of mRNAs as well. (Wagner, R. 1994. Nature 372:333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of a gene could be used in an antisense approach to inhibit translation of endogenous mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are typically less efficient inhibitors of translation but could also be used in accordance with the invention. Whether designed to hybridize to the 5', 3', or coding region of subject mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably less than about 100 and more preferably less than about 50, 25, 17 or 10 nucleotides in length.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors), or agents facilitating transport across the cell membrane (see, e.g., Letsinger *et al.*, 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre *et al.*, 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. WO 88/098 10, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. WO 89/10 134, published April 25, 1988), hybridization-triggered cleavage agents (See, e.g., Krol *et al.*, 1988, BioTechniques 6:958-976), or intercalating agents (See, e.g., Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to

another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-

5 chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxytriethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5-methoxycarboxymethyluracil, 10 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

15 The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

The antisense oligonucleotide can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, e.g., in Peny-  
20 O'Keefe *et al.* (1996) Proc. Natl. Acad. Sci. U.S.A. 93:14670 and in Eglom *et al.* (1993) Nature 365:566. One advantage of PNA oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a  
25 phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methyphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet a further embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each  
30 other (Gautier *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue *et al.*, 1987, Nucl. Acids Res. 15:6131-12148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, FEBS Lett. 215:327-330).

Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein *et al.* (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

While antisense nucleotides complementary to a coding region sequence can be used, those complementary to the transcribed untranslated region and to the region comprising the initiating methionine are most preferred.

The antisense molecules can be delivered to cells which express the target nucleic acid *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation on endogenous mRNAs. Therefore, a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pot II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous transcripts and thereby prevent translation of the target mRNA. For example, a vector can be introduced in vivo such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto *et al.*, 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner *et al.*, 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster



et al., 1982, Nature 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site; e.g., the choroid plexus or hypothalamus. Alternatively, viral vectors can be used which selectively infect the desired tissue (e.g., for brain, herpesvirus vectors may be used), in which case administration may be accomplished by another route (e.g., systemically).

In another aspect of the invention, ribozyme molecules designed to catalytically cleave target mRNA transcripts can be used to prevent translation of target mRNA and expression of a target protein (See, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver *et al.*, 1990, Science 247:1222-1225 and U.S. Patent No. 5,093,246). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature, 334:585-591. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224:574-578; Zaug and Cech, 1986, Science, 231:470-475; Zaug, et al., 1986, Nature, 324:429-433; published International patent application No. W088/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in a target gene.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells which express the target gene *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy

endogenous messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Antisense RNA, DNA, and ribozyme molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

#### VIII. Full-length cDNA Sequences of the Present Invention

The present invention also relates to full length cDNA sequences corresponding to one or more of the partial sequences of SEQ ID Nos. 1-4470. In particular the invention provides the full length cDNA sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. The full length sequences may be obtained as described above. These sequences are shown in Figure 2, and summarized below in Table 2. Also shown in Table 2 are the SEQ ID Nos and GenBank accession numbers for the polypeptides which are encoded by the full length cDNA sequences and which correspond to SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493.

cDNA SEQ ID NO.	Gene Name	GenBank Accession No.	Protein SEQ ID NO.	GenBank Accession No.
4472	Reg IV	NM 032044	4471	NP 114433
4474	XAG-2	NM 006408	4473	NP 006399

4476	SPARC/Osteonectin	NM 003118	4475	NP 003109
4478	GW112 protein	NM 006418	4477	NP 006409
4480	HSBP1	NM 001540	4479	NP 001531
4482	SKD1 Homolog	NP 004869	4481	NP 004860
4484	9-27	NM 003641	4483	NP 003632
4486	Defensin 5	NM 021010	4485	NP 066290
4488	p0071	NM 003628	4487	NP 003619
4490	UBE2I	NM 003345	4489	NP 003336
4492	Cytoplasmic dynein light chain	NM 003746	4491	NP 003737
4494	10Ckshs1	NM 001798	4493	NP 001789

#### IX. Polypeptides of the Present Invention

The present invention makes available isolated polypeptides which are isolated from, or otherwise substantially free of other cellular proteins, especially other signal transduction factors and/or transcription factors which may normally be associated with the polypeptide. Subject

5 polypeptides of the present invention include polypeptides encoded by the nucleic acids of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and

10 4494, or a sequence complementary thereto, or polypeptides encoded by genes of which a sequence in SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, or a sequence complementary thereto, is a fragment. In a preferred embodiment, polypeptides, useful in the present invention have the amino acid sequence of one or more of SEQ ID Nos. 4471,

15 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. Polypeptides of the present invention include those proteins which are differentially regulated in tumor cells, especially colon cancer-derived cell lines (relative to normal cells, e.g., normal colon tissue and

non-colon tissue). In a preferred embodiment the differentially regulated polypeptides are one or more of the polypeptides having the sequence set forth in SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. In preferred embodiments, the polypeptides are upregulated in tumor cells, especially colon cancer cancer-derived cell lines. In other embodiments, the polypeptides are downregulated in tumor cells, especially colon cancer-derived cell lines. Proteins which are upregulated, such as oncogenes, or downregulated, such as tumor suppressors, in aberrantly proliferating cells may be targets for diagnostic or therapeutic techniques. For example, upregulation of the *cdc2* gene induces mitosis. Overexpression of the *myt1* gene, a mitotic deactivator, negatively regulates the activity of *cdc2*. Aberrant proliferation may thus be induced either by upregulating *cdc2* or by downregulating *myt1*.

The term “substantially free of other cellular proteins” (also referred to herein as “contaminating proteins”) or “substantially pure or purified preparations” are defined as encompassing preparations of polypeptides having less than about 20% (by dry weight) contaminating protein, and preferably having less than about 5% contaminating protein.

Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned nucleic acid as described herein. Full length proteins or fragments corresponding to one or more particular motifs and/or domains or to arbitrary sizes, for example, at least about 5, 10, 25, 50, 75, or 100 amino acids in length are within the scope of the present invention.

For example, isolated polypeptides can be encoded by all or a portion of a nucleic acid sequence shown in any of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503 and most preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto. Isolated peptidyl portions of proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, a polypeptide of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., “authentic”) protein.

Another aspect of the present invention concerns recombinant forms of the subject proteins. Recombinant polypeptides preferred by the present invention, in addition to native proteins, as described above are encoded by a nucleic acid, which is at least 60%, more preferably at least 80%, and more preferably 85%, and more preferably 90%, and more preferably 95% identical to an amino acid sequence encoded by SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. Polypeptides which are encoded by a nucleic acid that is at least about 98-99% identical with the sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 are also within the scope of the invention. Also included in the present invention are peptide fragments comprising at least a portion of such a protein.

In a preferred embodiment, a polypeptide of the present invention is a mammalian polypeptide and even more preferably a human polypeptide. In particularly preferred embodiment, the polypeptide retains wild-type bioactivity. It will be understood that certain post-translational modifications, e.g., phosphorylation and the like, can increase the apparent molecular weight of the polypeptide relative to the unmodified polypeptide chain.

The present invention further pertains to recombinant forms of one of the subject polypeptides. Such recombinant polypeptides preferably are capable of functioning in one of either role of agonist or antagonist of at least one biological activity of a wild-type ("authentic") polypeptide of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of proteins, refers to both polypeptides having amino acid sequences which have arisen naturally, and also to mutational variants of human polypeptides which are derived, for example, by combinatorial mutagenesis.

In general, polypeptides referred to herein as having an activity (e.g., are "bioactive") of a protein are defined as polypeptides which include an amino acid sequence encoded by all or a portion of the nucleic acid sequences shown in one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and most preferably SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, or a sequence complementary thereto, and which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring protein. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally occurring form of a protein.

Assays for determining whether a compound, e.g, a protein or variant thereof, has one or more of the above biological activities are well known in the art. In certain embodiments, the polypeptides of the present invention have activities such as those outlined above.

In another embodiment, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable to produce an immunogenic fragment of a polypeptide (see, for example, EP Publication No: 0259149; and Evans *et al.* (1989) *Nature* 339:3 85; Huang *et al.* (1988) *J. Virol.* 62:3 855; and Schlienger *et al.*, (1992) *J. Virol.* 66:2). In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and, accordingly, can be used in the expression of the polypeptides of the present invention (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* (N.Y. John Wiley & Sons, 1991)). In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the expressed fusion protein by affinity chromatography using a Ni<sup>2+</sup> metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified protein (e.g., see Hochuli *et al.* (1987) *J. Chromatography* 411:177; and Janknecht *et al.* *PNAS* 88:8972).

Techniques for making fusion genes are known to those skilled in the art. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of nucleic acid fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive nucleic acid fragments which can subsequently be annealed to generate a chimeric nucleic acid sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992).

The present invention further pertains to methods of producing the subject polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. Suitable media for cell culture are well

known in the art. The recombinant polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptide. In a preferred  
5 embodiment, the recombinant polypeptide is a fusion protein containing a domain which facilitates its purification, such as GST fusion protein.

Moreover, it will be generally appreciated that, under certain circumstances, it may be advantageous to provide homologs of one of the subject polypeptides which function in a limited capacity as one of either an agonist (mimetic) or an antagonist, in order to promote or inhibit  
10 only a subset of the biological activities of the naturally occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of subject proteins.

Homologs of each of the subject polypeptide can be generated by mutagenesis, such as  
15 by discrete point mutation(s), or by truncation. For instance, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the polypeptide from which it was derived. Alternatively, antagonistic forms of the polypeptide can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to a receptor.

20 The recombinant polypeptides of the present invention also include homologs of the wild-type proteins, such as versions of those proteins which are resistant to proteolytic cleavage, for example, due to mutations which alter ubiquitination or other enzymatic targeting associated with the protein.

Polypeptides may also be chemically modified to create derivatives by forming covalent  
25 or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of proteins can be prepared by linking the chemical moieties to functional groups on amino acid sidechains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

Modification of the structure of the subject polypeptides can be for such purposes as  
30 enhancing therapeutic or prophylactic efficacy, stability (e.g., *ex vivo* shelf life and resistance to proteolytic degradation), or post-translational modifications (e.g., to alter phosphorylation pattern of protein). Such modified peptides, when designed to retain at least one activity of the

naturally occurring form of the protein, or to produce specific antagonists thereof, are considered functional equivalents of the polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid substitution, deletion, or addition. The substitutional variant may be a substituted conserved amino acid or a substituted non-conserved amino acid.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e., isosteric and/or isoelectric mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tiyptophan; (5) amide = asparagine, glutamine; and (6) sulfur -containing = cysteine and methionine. (see, for example, *Biochemistry*, 2 ed., Ed. by L. Stryer, WH Freeman and Co.: 1981). Whether a change in the amino acid sequence of a peptide results in a functional homolog (e.g., functional in the sense that the resulting polypeptide mimics or antagonizes the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response.

Polypeptides in which more than one replacement has taken place can readily be tested in the same manner. The variant may be designed so as to retain biological activity of a particular region of the protein. In a non-limiting example, Osawa et al., 1994, *Biochemistry and Molecular International* 34:1003-1009, discusses the actin binding region of a protein from several different species. The actin binding regions of the these species are considered homologous based on the fact that they have amino acids that fall within "homologous residue groups." Homologous residues are judged according to the following groups (using single letter amino acid designations): STAG; ILVMF; HRK; DEQN; and FYW. For example, an S, a T, an A or a G can be in a position and the function (in this case actin binding) is retained.



Additional guidance on amino acid substitution is available from studies of protein evolution. Go et al., 1980, *Int. J. Peptide Protein Res.* 15: 211-224, classified amino acid residue sites as interior or exterior depending on their accessibility. More frequent substitution on exterior sites was confirmed to be general in eight sets of homologous protein families regardless of their biological functions and the presence or absence of a prosthetic group. Virtually all types of amino acid residues had higher mutabilities on the exterior than in the interior. No correlation between mutability and polarity was observed of amino acid residues in the interior and exterior, respectively. Amino acid residues were classified into one of three groups depending on their polarity: polar (Arg, Lys, His, Gln, Asn, Asp, and Glu); weak polar (Ala, Pro, Gly, Thr, and Ser), and nonpolar (Cys, Val, Met, Ile, Leu, Phe, Tyr, and Trp). Amino acid replacements during protein evolution were very conservative: 88% and 76% of them in the interior or exterior, respectively, were within the same group of the three. Intergroup replacements are such that weak polar residues are replaced more often by nonpolar residues in the interior and more often by polar residues on the exterior.

Querol et al., 1996, *Prot. Eng.* 9:265-271, provides general rules for amino acid substitutions to enhance protein thermostability. New glycosylation sites can be introduced as discussed in Olsen and Thomsen, 1991, *J. Gen. Microbiol.* 137 :579-585. An additional disulfide bridge can be introduced, as discussed by Perry and Wetzel, 1984, *Science* 226:555-557; Pantoliano et al., 1987, *Biochemistry* 26:2077-2082; Matsumura et al., 1989, *Nature* 342:291-293; Nishikawa et al., 1990, *Protein Eng.* 3:443-448; Takagi et al., 1990, *J. Biol. Chem.*, 265:6874-6878; Clarke et al., 1993, *Biochemistry* 32:4322-4329; and Wakarchuk et al., 1994, *Protein Eng.* 7:1379-1386.

An additional metal binding site can be introduced, according to Toma et al., 1991, *Biochemistry* 30:97-106, and Haezebrouck et al., 1993, *Protein Eng.* 6:643-649. Substitutions with prolines in loops can be made according to Masul et al., 1994, *Appl Env. Microbiol.* 60:3579-3584; and Hardy et al., *FEBS Lett.* 317:89-92.

Cysteine-depleted muteins are considered variants within the scope of the invention. These variants can be constructed according to methods disclosed in U.S. Patent No. 4,959,314, which discloses how to substitute other amino acids for cysteines, and how to determine biological activity and effect of the substitution. Such methods are suitable for proteins according to this invention that have cysteine residues suitable for such substitutions, for example to eliminate disulfide bond formation.

To learn the identity and function of the gene that correlates with an nucleic acid, the nucleic acids or corresponding amino acid sequences can be screened against profiles of protein families. Such profiles focus on common structural motifs among proteins of each family. Publicly available profiles are described above.

5           In comparing a new nucleic acid with known sequences, several alignment tools are available. Examples include PileUp, which creates a multiple sequence alignment, and is described in Feng *et al.*, *J. Mol. Evol.* (1987) 25:35 1-360. Another method, GAP, uses the alignment method of Needleman *et al.*, *J. Mol. Biol.* (1970) 48:443-453. GAP is best suited for global alignment of sequences. A third method, BestFit, functions by inserting gaps to maximize  
10   the number of matches using the local homology algorithm of Smith and Waterman, *Adv. Appl. Math.* (1981) 2:482-489.

#### X.     Diagnostic & Prognostic Assays and Drug Screening Methods

          The present invention provides method for determining whether a subject is at risk for developing a disease or condition characterized by unwanted cell proliferation by detecting the  
15   disclosed biomarkers, i.e., the present nucleic acids (SEQ ID Nos: 1-4494) and/or polypeptide markers (preferably SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493) for colon cancer encoded thereby.

          In clinical applications, human tissue samples can be screened for the presence and/or absence of the biomarkers identified herein. Such samples could consist of needle biopsy cores,  
20   surgical resection samples, lymph node tissue, or serum. For example, these methods include obtaining a biopsy, which is optionally fractionated by cryostat sectioning to enrich tumor cells to about 80% of the total cell population. In certain embodiments, nucleic acids extracted from these samples may be amplified using techniques well known in the art. The levels of selected markers detected would be compared with statistically valid groups of metastatic, non-metastatic  
25   malignant, benign, or normal colon tissue samples.

          In one embodiment, the diagnostic method comprises determining whether a subject has an abnormal mRNA and/or protein level of the disclosed markers, such as by Northern blot analysis, reverse transcription-polymerase chain reaction (RT-PCR), in situ hybridization, immunoprecipitation, Western blot hybridization, or immunohistochemistry. According to the  
30   method, cells are obtained from a subject and the levels of the disclosed biomarkers, protein or mRNA level, is determined and compared to the level of these markers in a healthy subject. An

abnormal level of the biomarker polypeptide or mRNA levels is likely to be indicative of cancer such as colon cancer.

Accordingly, in one aspect, the invention provides probes and primers that are specific to the unique nucleic acid markers disclosed herein. Accordingly, the nucleic acid probes comprise  
5 a nucleotide sequence at least 10 nucleotides in length, preferably at least 15 nucleotides, more preferably, 25 nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all of the coding sequence which is complementary to a portion of the coding sequence of a marker nucleic acid sequence, which nucleic acid sequence is represented by SEQ ID Nos: 1-4494 or a sequence complementary thereto.

10 In one embodiment, the method comprises using a nucleic acid probe to determine the presence of cancerous cells in a tissue from a patient. Specifically, the method comprises:

1. providing a nucleic acid probe comprising a nucleotide sequence at least 10 nucleotides in length, preferably at least 15 nucleotides, more preferably, 25 nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all  
15 of the coding sequence which is complementary to a portion of the coding sequence of a nucleic acid sequence represented by SEQ ID Nos: 1-4494 or a sequence complementary thereto and is differentially expressed in tumors cells, such as colon cancer cells;
2. obtaining a tissue sample from a patient potentially comprising cancerous cells;
- 20 3. providing a second tissue sample containing cells substantially all of which are non-cancerous;
4. contacting the nucleic acid probe under stringent conditions with RNA of each of said first and second tissue samples (e.g., in a Northern blot or in situ hybridization assay); and
- 25 5. comparing (a) the amount of hybridization of the probe with RNA of the first tissue sample, with (b) the amount of hybridization of the probe with RNA of the second tissue sample; wherein a statistically significant difference in the amount of hybridization with the RNA of the first tissue sample as compared to the amount of hybridization with the RNA of the second tissue sample is indicative of  
30 the presence of cancerous cells in the first tissue sample.

In one aspect, the method comprises *in situ* hybridization with a probe derived from a given marker nucleic acid sequence, which nucleic acid sequence is represented by SEQ ID Nos: 1-4494 or a sequence complementary thereto. The method comprises contacting the labeled hybridization probe with a sample of a given type of tissue potentially containing cancerous or pre-cancerous cells as well as normal cells, and determining whether the probe labels some cells of the given tissue type to a degree significantly different (e.g., by at least a factor of two, or at least a factor of five, or at least a factor of twenty, or at least a factor of fifty) than the degree to which it labels other cells of the same tissue type.

Also within the invention is a method of determining the phenotype of a test cell from a given human tissue, e.g., whether the cell is (a) normal, or (b) cancerous or precancerous, by contacting the mRNA of a test cell with a nucleic acid probe at least 12 nucleotides in length, preferably at least 15 nucleotides, more preferably at least 25 nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all of a sequence which is complementary to a portion of the coding sequence of a nucleic acid sequence represented by SEQ ID Nos: 1-4494 or a sequence complementary thereto, and which is differentially expressed in tumor cells as compared to normal cells of the given tissue type; and determining the approximate amount of hybridization of the probe to the mRNA, an amount of hybridization either more or less than that seen with the mRNA of a normal cell of that tissue type being indicative that the test cell is cancerous or pre-cancerous.

Alternatively, the above diagnostic assays may be carried out using antibodies to detect the protein product encoded by the marker nucleic acid sequence, which nucleic acid sequence is represented by SEQ ID Nos: 1-4494 or a sequence complementary thereto. Accordingly, in one embodiment, the assay would include contacting the proteins of the test cell with an antibody specific for the gene product of a nucleic acid represented by SEQ ID Nos: 1-4494, preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, the marker nucleic acid being one which is expressed at a given control level in normal cells of the same tissue type as the test cell, and determining the approximate amount of immunocomplex formation by the antibody and the proteins of the test cell, wherein a statistically significant difference in the amount of the immunocomplex formed with the proteins of a test cell as compared to a normal cell of the same tissue type is an indication that the test cell is cancerous or pre-cancerous. Preferably, the antibody is specific for one of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493.

The method for producing polyclonal and/or monoclonal antibodies which specifically bind to polypeptides useful in the present invention is known to those of skill in the art and can be found in, for example Dymecki et al., 1992, J. Biol. Chem., 267:4815; Boersma & Van Leeuwen, 1994, J. Neurosci. Methods, 51:317; Green et al., 1982, Cell, 28:477; and Arnheiter et al., 1981, Nature, 294:278.

Another such method includes the steps of: providing an antibody specific for the gene product of a marker nucleic acid sequence represented by SEQ ID Nos 1-4494, the gene product being present in cancerous tissue of a given tissue type (e.g., colon tissue) at a level more or less than the level of the gene product in non-cancerous tissue of the same tissue type; obtaining from a patient a first sample of tissue of the given tissue type, which sample potentially includes cancerous cells; providing a second sample of tissue of the same tissue type (which may be from the same patient or from a normal control, e.g. another individual or cultured cells), this second sample containing normal cells and essentially no cancerous cells; contacting the antibody with protein (which may be partially purified, in lysed but unfractionated cells, or in situ) of the first and second samples under conditions permitting immunocomplex formation between the antibody and the marker nucleic acid sequence product present in the samples; and comparing (a) the amount of immunocomplex formation in the first sample, with (b) the amount of immunocomplex formation in the second sample, wherein a statistically significant difference in the amount of immunocomplex formation in the first sample less as compared to the amount of immunocomplex formation in the second sample is indicative of the presence of cancerous cells in the first sample of tissue.

The subject invention further provides a method of determining whether a cell sample obtained from a subject possesses an abnormal amount of marker polypeptide which comprises (a) obtaining a cell sample from the subject, (b) quantitatively determining the amount of the marker polypeptide in the sample so obtained, and (c) comparing the amount of the marker polypeptide so determined with a known standard, so as to thereby determine whether the cell sample obtained from the subject possesses an abnormal amount of the marker polypeptide. Such marker polypeptides may be detected by immunohistochemical assays, dot-blot assays, ELISA and the like.

Immunoassays are commonly used to quantitate the levels of proteins in cell samples, and many other immunoassay techniques are known in the art. The invention is not limited to a particular assay procedure, and therefore is intended to include both homogeneous and heterogeneous procedures. Exemplary immunoassays which can be conducted according to the

invention include fluorescence polarization immunoassay (FPIA), fluorescence immunoassay (FIA), enzyme immunoassay (EIA), nephelometric inhibition immunoassay (NIA), enzyme linked immunosorbent assay (ELISA), and radioimmunoassay (RIA). An indicator moiety, or label group, can be attached to the subject antibodies and is selected so as to meet the needs of various uses of the method which are often dictated by the availability of assay equipment and compatible immunoassay procedures. General techniques to be used in performing the various immunoassays noted above are known to those of ordinary skill in the art.

In another embodiment, the level of the encoded product, i.e., the product encoded by SEQ ID Nos 1-4494 or a sequence complementary thereto, or alternatively the level of the polypeptide of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, in a biological fluid (e.g., blood or urine) of a patient may be determined as a way of monitoring the level of expression of the marker nucleic acid sequence in cells of that patient. Such a method would include the steps of obtaining a sample of a biological fluid from the patient, contacting the sample (or proteins from the sample) with an antibody specific for a encoded marker polypeptide, and determining the amount of immune complex formation by the antibody, with the amount of immune complex formation being indicative of the level of the marker encoded product in the sample. This determination is particularly instructive when compared to the amount of immune complex formation by the same antibody in a control sample taken from a normal individual or in one or more samples previously or subsequently obtained from the same person.

In another embodiment, the method can be used to determine the amount of marker polypeptide present in a cell, which in turn can be correlated with progression of a hyperproliferative disorder, e.g., colon cancer. The level of the marker polypeptide can be used predictively to evaluate whether a sample of cells contains cells which are, or are predisposed towards becoming, transformed cells. Moreover, the subject method can be used to assess the phenotype of cells which are known to be transformed, the phenotyping results being useful in planning a particular therapeutic regimen. For instance, very high levels of the marker polypeptide in sample cells is a powerful diagnostic and prognostic marker for a cancer, such as colon cancer. The observation of marker polypeptide level can be utilized in decisions regarding, e.g., the use of more aggressive therapies.

As set out above, one aspect of the present invention relates to diagnostic assays for determining, in the context of cells isolated from a patient, if the level of a marker polypeptide is significantly reduced in the sample cells. The term "significantly reduced" refers to a cell

phenotype wherein the cell possesses a reduced cellular amount of the marker polypeptide relative to a normal cell of similar tissue origin. For example, a cell may have less than about 50%, 25%, 10%, or 5% of the marker polypeptide that a normal control cell. In particular, the assay evaluates the level of marker polypeptide in the test cells, and, preferably, compares the measured level with marker polypeptide detected in at least one control cell, e.g., a normal cell and/or a transformed cell of known phenotype.

Of particular importance to the subject invention is the ability to quantitate the level of marker polypeptide as determined by the number of cells associated with a normal or abnormal marker polypeptide level. The number of cells with a particular marker polypeptide phenotype may then be correlated with patient prognosis. In one embodiment of the invention, the marker polypeptide phenotype of the lesion is determined as a percentage of cells in a biopsy which are found to have abnormally high/low levels of the marker polypeptide. Such expression may be detected by immunohistochemical assays, dot-blot assays, ELISA and the like.

Where tissue samples are employed, immunohistochemical staining may be used to determine the number of cells having the marker polypeptide phenotype. For such staining, a multiblock of tissue is taken from the biopsy or other tissue sample and subjected to proteolytic hydrolysis, employing such agents as protease K or pepsin. In certain embodiments, it may be desirable to isolate a nuclear fraction from the sample cells and detect the level of the marker polypeptide in the nuclear fraction.

The tissue samples are fixed by treatment with a reagent such as formalin, glutaraldehyde, methanol, or the like. The samples are then incubated with an antibody, preferably a monoclonal antibody, with binding specificity for the marker polypeptides. This antibody may be conjugated to a label for subsequent detection of binding. Samples are incubated for a time sufficient for formation of the immunocomplexes. Binding of the antibody is then detected by virtue of a label conjugated to this antibody. Where the antibody is unlabeled, a second labeled antibody may be employed, e.g., which is specific for the isotype of the anti-marker polypeptide antibody. Examples of labels which may be employed include radionuclides, fluorescers, chemiluminescers, enzymes and the like.

Where enzymes are employed, the substrate for the enzyme may be added to the samples to provide a colored or fluorescent product. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-enzyme conjugates are readily produced by techniques known to those skilled in the art.

In one embodiment, the assay is performed as a dot blot assay. The dot blot assay finds particular application where tissue samples are employed as it allows determination of the average amount of the marker polypeptide associated with a single cell by correlating the amount of marker polypeptide in a cell-free extract produced from a predetermined number of cells.

5           It is well established in the cancer literature that tumor cells of the same type (e.g., breast and/or colon tumor cells) may not show uniformly increased expression of individual oncogenes or uniformly decreased expression of individual tumor suppressor genes. There may also be varying levels of expression of a given marker gene even between cells of a given type of cancer, further emphasizing the need for reliance on a battery of tests rather than a single test.  
10       Accordingly, in one aspect, the invention provides for a battery of tests utilizing a number of probes of the invention, in order to improve the reliability and/or accuracy of the diagnostic test.

          In one embodiment, the present invention also provides a method wherein nucleic acid probes are immobilized on a DNA chip in an organized array. Oligonucleotides can be bound to a solid support by a variety of processes, including lithography. For example a chip can hold up  
15       to 250,000 oligonucleotides (GeneChip, Affymetrix). These nucleic acid probes comprise a nucleotide sequence at least about 12 nucleotides in length, preferably at least about 15 nucleotides, more preferably at least about 25 nucleotides, and most preferably at least about 40 nucleotides, and up to all or nearly all of a sequence which is complementary to a portion of the coding sequence of a marker nucleic acid sequence represented by SEQ ID Nos: 1-4494 and is  
20       differentially expressed in tumor cells, such as colon cancer cells. The present invention provides significant advantages over the available tests for various cancers, such as colon cancer, because it increases the reliability of the test by providing an array of nucleic acid markers on a single chip.

          The method includes obtaining a biopsy, which is optionally fractionated by cryostat  
25       sectioning to enrich tumor cells to about 80% of the total cell population. The DNA or RNA is then extracted, amplified, and analyzed with a DNA chip to determine the presence of absence of the marker nucleic acid sequences.

          In one embodiment, the nucleic acid probes are spotted onto a substrate in a two-dimensional matrix or array. Samples of nucleic acids can be labeled and then hybridized to the  
30       probes. Double-stranded nucleic acids, comprising the labeled sample nucleic acids bound to probe nucleic acids, can be detected once the unbound portion of the sample is washed away.



The probe nucleic acids can be spotted on substrates including glass, nitrocellulose, etc. The probes can be bound to the substrate by either covalent bonds or by non-specific interactions, such as hydrophobic interactions. The sample nucleic acids can be labeled using radioactive labels, fluorophores, chromophores, etc.

5           Techniques for constructing arrays and methods of using these arrays are described, for example, in EP No. 0 799 897; PCT No. WO 97/292 12; PCT No. WO 97127317; EP No. 0 785 280; PCT No. WO 97/02357; U.S. Pat. No. 5,593,839; U.S. Pat. No. 5,578,832; EP No. 0 728 520; U.S. Pat. No. 5,599,695; EP No. 0 721 016; U.S. Pat. No. 5,556,752; PCT No. WO 95/22058; and U.S. Pat. No. 5,631,734.

10           Further, arrays can be used to examine differential expression of genes and can be used to determine gene function. For example, arrays of the instant nucleic acid sequences can be used to determine if any of the nucleic acid sequences are differentially expressed between normal cells and cancer cells, for example. High expression of a particular message in a cancer cell, which is not observed in a corresponding normal cell, can indicate a cancer specific protein.

15           In one embodiment nucleic acid molecules useful in the present invention, such as those of SEQ ID Nos 1-4494, preferably those of SEQ ID Nos 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, may be used to generate macroarrays on a solid surface such as a membrane such that the arrayed nucleic acid molecules can be used to determine if any of the nucleic acids are differentially expressed between normal cells or tissue and cancerous  
20           cells or tissue. In one embodiment, the nucleic acid molecules of the invention are either cDNA or may be used to generate cDNA molecules to be subsequently amplified by PCR and spotted on nylon membranes. The membranes are then reacted with radiolabeled target nucleic acid molecules obtained from equivalent samples of cancerous and normal tissue or cells. Methods of cDNA generation and macroarray preparation are known to those of skill in the art and may be  
25           found, for example in Bertucci et al., 1999 *Hum. Mol. Genet.* 8:2129; Nguyen et al., 1995, *Genomics*, 29: 207; Zhao et al., *Gene*, 156:207; Gress et al., 1992, *Mammalian Genome*, 3:609; Zhumabayeva et al., 2001, *Biotechniques*, 30:158; and Lennon et al., 1991, *Trends Genet.* 7:314.

          In yet another embodiment, the invention contemplates using a panel of antibodies which are generated against the marker polypeptides of this invention, which polypeptides are encoded  
30           by one or more of SEQ ID Nos: 1-4494, preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494. Preferably, the antibodies are generated against one or more polypeptides having the sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. Such a panel of antibodies may be used as a

reliable diagnostic probe for colon cancer. The assay of the present invention comprises contacting a biopsy sample containing cells, e.g., colon cells, with a panel of antibodies to one or more of the encoded products to determine the presence or absence of the marker polypeptides.

5 The diagnostic methods of the subject invention may also be employed as follow-up to treatment, e.g., quantitation of the level of marker polypeptides may be indicative of the effectiveness of current or previously employed cancer therapies as well as the effect of these therapies upon patient prognosis.

Accordingly, the present invention makes available diagnostic assays and reagents for detecting gain and/or loss of marker polypeptides from a cell in order to aid in the diagnosis and  
10 phenotyping of proliferative disorders arising from, for example, tumorigenic transformation of cells.

The diagnostic assays described above can be adapted to be used as prognostic assays, as well. Such an application takes advantage of the sensitivity of the assays of the invention to events which take place at characteristic stages in the progression of a tumor. For example, a  
15 given marker gene may be up- or downregulated at a very early stage, perhaps before the cell is irreversibly committed to developing into a malignancy, while another marker gene may be characteristically up or down regulated only at a much later stage. Such a method could involve the steps of contacting the mRNA of a test cell with a nucleic acid probe derived from a given marker nucleic acid which is expressed at different characteristic levels in cancerous or  
20 precancerous cells at different stages of tumor progression, and determining the approximate amount of hybridization of the probe to the mRNA of the cell, such amount being an indication of the level of expression of the gene in the cell, and thus an indication of the stage of tumor progression of the cell; alternatively, the assay can be carried out with an antibody specific for the gene product of the given marker nucleic acid, contacted with the proteins of the test cell. A  
25 battery of such tests will disclose not only the existence and location of a tumor, but also will allow the clinician to select the mode of treatment most appropriate for the tumor, and to predict the likelihood of success of that treatment.

The methods of the invention can also be used to follow the clinical course of a tumor. For example, the assay of the invention can be applied to a tissue sample from a patient;  
30 following treatment of the patient for the cancer, another tissue sample is taken and the test repeated. Successful treatment will result in either removal of all cells which demonstrate differential expression characteristic of the cancerous or precancerous cells, or a substantial

increase in expression of the gene in those cells, perhaps approaching or even surpassing normal levels.

In yet another embodiment, the invention provides methods for determining whether a subject is at risk for developing a disease, such as a predisposition to develop cancer, for example colon cancer, associated with an aberrant activity of any one of the polypeptides encoded by nucleic acids of SEQ ID Nos: 1-4494, preferably, any one of the polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, wherein the aberrant activity of the polypeptide is characterized by detecting the presence or absence of a genetic lesion characterized by at least one of (i) an alteration affecting the integrity of a gene encoding a marker polypeptides, or (ii) the mis-expression of the encoding nucleic acid. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from the nucleic acid sequence, (ii) an addition of one or more nucleotides to the nucleic acid sequence, (iii) a substitution of one or more nucleotides of the nucleic acid sequence, (iv) a gross chromosomal rearrangement of the nucleic acid sequence, (v) a gross alteration in the level of a messenger RNA transcript of the nucleic acid sequence, (vi) aberrant modification of the nucleic acid sequence, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene, (viii) a non-wild type level of the marker polypeptide, (ix) allelic loss of the gene, and/or (x) inappropriate post-translational modification of the marker polypeptide.

The present invention provides assay techniques for detecting lesions in the encoding nucleic acid sequence. These methods include, but are not limited to, methods involving sequence analysis, Southern blot hybridization, restriction enzyme site mapping, and methods involving detection of absence of nucleotide pairing between the nucleic acid to be analyzed and a probe.

Specific diseases or disorders, e.g., genetic diseases or disorders, are associated with specific allelic variants of polymorphic regions of certain genes, which do not necessarily encode a mutated protein. Thus, the presence of a specific allelic variant of a polymorphic region of a gene in a subject can render the subject susceptible to developing a specific disease or disorder. Polymorphic regions in genes, can be identified, by determining the nucleotide sequence of genes in populations of individuals. If a polymorphic region is identified, then the link with a specific disease can be determined by studying specific populations of individuals, e.g., individuals which developed a specific disease, such as colon cancer. A polymorphic region can

be located in any region of a gene, e.g., exons, in coding or non coding regions of exons, introns, and promoter region.

In an exemplary embodiment, there is provided a nucleic acid composition comprising a nucleic acid probe including a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence of a gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is contacted with the nucleic acid of the sample, and the hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect lesions or allelic variants at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

A preferred detection method is allele specific hybridization using probes overlapping the mutation or polymorphic site and having about 5, 10, 20, 25, or 30 nucleotides around the mutation or polymorphic region. In a preferred embodiment of the invention, several probes capable of hybridizing specifically to allelic variants are attached to a solid phase support, e.g., a "chip". Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described e.g., in Cronin et al. (1996) *Human Mutation* 7:244. In one embodiment, a chip comprises all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted with a test nucleic acid and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment.

In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligase chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the gene (see Abravaya *et al.* (1995) *Nuc Acid Res* 23:675-682). In a merely illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the nucleic acid sample with one or more primers which specifically hybridize to a nucleic acid sequence under conditions such that hybridization and amplification of the nucleic acid (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR

and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional amplification system  
5 (Kwoh, D.Y. *et al.*, 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.*, 1988, Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

10 In a preferred embodiment of the subject assay, mutations in, or allelic variants, of a gene from a sample cell are identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531)  
15 can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

Another aspect of the invention is directed to the identification of agents capable of modulating the differentiation and proliferation of cells characterized by aberrant proliferation. In this regard, the invention provides assays for determining compounds that modulate the  
20 expression of the marker nucleic acids (SEQ ID Nos: 1-4494, preferably SEQ ID Nos 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494) and/or alter for example, inhibit the bioactivity of the encoded polypeptide such as those of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493.

Several *in vivo* methods can be used to identify compounds that modulate expression of  
25 the marker nucleic acids (SEQ ID Nos: 1-4494) and/or alter for example, inhibit the bioactivity of the encoded polypeptide.

Drug screening is performed by adding a test compound to a sample of cells, and monitoring the effect. A parallel sample which does not receive the test compound is also monitored as a control. The treated and untreated cells are then compared by any suitable  
30 phenotypic criteria, including but not limited to microscopic analysis, viability testing, ability to replicate, histological examination, the level of a particular RNA or polypeptide associated with the cells, the level of enzymatic activity expressed by the cells or cell lysates, and the ability of

the cells to interact with other cells or compounds. Differences between treated and untreated cells indicates effects attributable to the test compound.

Desirable effects of a test compound include an effect on any phenotype that was conferred by the cancer-associated marker nucleic acid sequence. Examples include a test compound that limits the overabundance of mRNA, limits production of the encoded protein, or limits the functional effect of the protein. The effect of the test compound would be apparent when comparing results between treated and untreated cells.

The invention thus also encompasses methods of screening for agents which inhibit expression of the nucleic acid markers (SEQ ID Nos: 1-4494, preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494) *in vitro*, comprising exposing a cell or tissue in which the marker nucleic acid mRNA is detectable in cultured cells to an agent in order to determine whether the agent is capable of inhibiting production of the mRNA; and determining the level of mRNA in the exposed cells or tissue, wherein a decrease in the level of the mRNA after exposure of the cell line to the agent is indicative of inhibition of the marker nucleic acid mRNA production.

Alternatively, the screening method may include *in vitro* screening of a cell or tissue in which marker protein is detectable in cultured cells to an agent suspected of inhibiting production of the marker protein; and determining the level of the marker protein in the cells or tissue, wherein a decrease in the level of marker protein after exposure of the cells or tissue to the agent is indicative of inhibition of marker protein production.

The invention also encompasses *in vivo* methods of screening for agents which inhibit expression of the marker nucleic acids, comprising exposing a mammal having tumor cells in which marker mRNA or protein is detectable to an agent suspected of inhibiting production of marker mRNA or protein; and determining the level of marker mRNA or protein in tumor cells of the exposed mammal. A decrease in the level of marker mRNA or protein after exposure of the mammal to the agent is indicative of inhibition of marker nucleic acid expression.

Accordingly, the invention provides a method comprising incubating a cell expressing the marker nucleic acids (SEQ ID Nos: 1-4494) with a test compound and measuring the mRNA or protein level. The invention further provides a method for quantitatively determining the level of expression of the marker nucleic acids in a cell population, and a method for determining whether an agent is capable of increasing or decreasing the level of expression of the marker nucleic acids in a cell population. The method for determining whether an agent is capable of

increasing or decreasing the level of expression of the marker nucleic acids in a cell population comprises the steps of (a) preparing cell extracts from control and agent-treated cell populations, (b) isolating the marker polypeptides from the cell extracts, (c) quantifying (e.g., in parallel) the amount of an immunocomplex formed between the marker polypeptide and an antibody specific to said polypeptide. The marker polypeptides of this invention may also be quantified by assaying for its bioactivity. Agents that induce increased the marker nucleic acid expression may be identified by their ability to increase the amount of immunocomplex formed in the treated cell as compared with the amount of the immunocomplex formed in the control cell. In a similar manner, agents that decrease expression of the marker nucleic acid may be identified by their ability to decrease the amount of the immunocomplex formed in the treated cell extract as compared to the control cell.

mRNA levels can be determined by Northern blot hybridization. mRNA levels can also be determined by methods involving PCR. Other sensitive methods for measuring mRNA, which can be used in high throughput assays, e.g., a method using a DELFIA endpoint detection and quantification method, are described, e.g., in Webb and Hurskainen (1996) *Journal of Biomolecular Screening* 1:119. Marker protein levels can be determined by immunoprecipitations or immunohistochemistry using an antibody that specifically recognizes the protein product encoded by SEQ ID Nos: 1- 4494, and preferably one or more of the proteins having the sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493.

Agents that are identified as active in the drug screening assay are candidates to be tested for their capacity to block cell proliferation activity. These agents would be useful for treating a disorder involving aberrant growth of cells, especially colon cells.

A variety of assay formats will suffice and, in light of the present disclosure, those not expressly described herein will nevertheless be comprehended by one of ordinary skill in the art. For instance, the assay can be generated in many different formats, and include assays based on cell-free systems, e.g., purified proteins or cell lysates, as well as cell-based assays which utilize intact cells.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays of the present invention which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins or with lysates, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively

easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the *in vitro* system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or changes in enzymatic properties of the molecular target.

A. Use of Nucleic Acids as Probes in Mapping and in Tissue Profiling Probes

Polynucleotide probes as described above, e g , comprising at least 12 contiguous nucleotides selected from the nucleotide SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, are used for a variety of purposes, including identification of human chromosomes and determining transcription levels. Additional disclosure about preferred regions of the nucleic acid sequences is found in the accompanying tables.

The nucleotide probes are labeled, for example, with a radioactive, fluorescent, biotinylated, or chemiluminescent label, and detected by well known methods appropriate for the particular label selected. Protocols for hybridizing nucleotide probes to preparations of metaphase chromosomes are also well known in the art. A nucleotide probe will hybridize specifically to nucleotide sequences in the chromosome preparations which are complementary to the nucleotide sequence of the probe. A probe that hybridizes specifically to a nucleic acid should provide a detection signal at least 5-, 10-, or 20-fold higher than the background hybridization provided with other unrelated sequences.

In a non-limiting example, commercial programs are available for identifying regions of chromosomes commonly associated with disease, such as cancer. Nucleic acids of the invention can be used to probe these regions. For example, if, through profile searching, a nucleic acid is identified as corresponding to a gene encoding a kinase, its ability to bind to a cancer-related chromosomal region will suggest its role as a kinase in one or more stages of tumor cell development/growth. Although some experimentation would be required to elucidate the role, the nucleic acid constitutes a new material for isolating a specific protein that has potential for developing a cancer diagnostic or therapeutic.

Nucleotide probes are used to detect expression of a gene corresponding to the nucleic acid. For example, in Northern blots, mRNA is separated electrophoretically and contacted with



a probe. A probe is detected as hybridizing to an mRNA species of a particular size. The amount of hybridization is quantitated to determine relative amounts of expression, for example under a particular condition. Probes are also used to detect products of amplification by polymerase chain reaction. The products of the reaction are hybridized to the probe and hybrids are detected.

- 5 Probes are used for *in situ* hybridization to cells to detect expression. Probes can also be used *in vivo* for diagnostic detection of hybridizing sequences. Probes are typically labeled with a radioactive isotope. Other types of detectable labels may be used such as chromophores, fluorophores, and enzymes.

Expression of specific mRNA can vary in different cell types and can be tissue specific.

- 10 This variation of mRNA levels in different cell types can be exploited with nucleic acid probe assays to determine tissue types. For example, PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes substantially identical or complementary to nucleic acids of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more  
15 preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, can determine the presence or absence of target cDNA or mRNA.

Examples of a nucleotide hybridization assay are described in Urdea *et al.*, PCT W092/02526 and Urdea *et al.*, U.S. Patent No. 5,124,246, both incorporated herein by reference.

- 20 The references describe an example of a sandwich nucleotide hybridization assay.

Alternatively, the Polymerase Chain Reaction (PCR) is another means for detecting small amounts of target nucleic acids, as described in Mullis *et al.*, *Met/i. Enzymol.* (1987) 155:335-350; U.S. Patent No. 4,683,195; and U.S. Patent No. 4,683,202, all incorporated herein by reference. Two primer polynucleotides hybridize with the target nucleic acids and  
25 are used to prime the reaction. The primers may be composed of sequence within or 3' and 5' to the polynucleotides of the Sequence Listing. Alternatively, if the primers are 3' and 5' to these polynucleotides, they need not hybridize to them or the complements. A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a large amount of target nucleic acids is generated by the  
30 polymerase, it is detected by methods such as Southern blots. When using the Southern blot method, the labeled probe will hybridize to a polynucleotide of the Sequence Listing or complement.

Furthermore, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1989). mRNA or cDNA generated from mRNA using a polymerase enzyme can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labeled probe and then washed to remove any unhybridized probe. Next, the duplexes containing the labeled probe are detected. Typically, the probe is labeled with radioactivity.

### Mapping

Nucleic acids of the present invention are used to identify a chromosome on which the corresponding gene resides. Using fluorescence *in situ* hybridization (FISH) on normal metaphase spreads, comparative genomic hybridization allows total genome assessment of changes in relative copy number of DNA sequences. See Schwartz and Samad, *Current Opinions in Biotechnology* (1994) 8:70-74; Kallioniemi *et al.*, *Seminars in Cancer Biology* (1993) 4:41-46; Valdes and Tagle, *Methods in Molecular Biology* (1997) 68:1, Boultonwood, ed., Human Press, Totowa, NJ.

Preparations of human metaphase chromosomes are prepared using standard cytogenetic techniques from human primary tissues or cell lines. Nucleotide probes comprising at least 12 contiguous nucleotides selected from the nucleotide sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, are used to identify the corresponding chromosome. The nucleotide probes are labeled, for example, with a radioactive, fluorescent, biotinylated, or chemiluminescent label, and detected by well known methods appropriate for the particular label selected. Protocols for hybridizing nucleotide probes to preparations of metaphase chromosomes are also well known in the art. A nucleotide probe will hybridize specifically to nucleotide sequences in the chromosome preparations that are complementary to the nucleotide sequence of the probe. A probe that hybridizes specifically to a target gene provides a detection signal at least 5-, 10-, or 20-fold higher than the background hybridization provided with unrelated coding sequences.

Nucleic acids are mapped to particular chromosomes using, for example, radiation hybrids or chromosome-specific hybrid panels. See Leach *et al.*, *Advances in Genetics*, (1995) 33:63-99; Walter *et al.*, *Nature Genetics* (1994) 7:22-28; Walter and Goodfellow, *Trends in*

*Genetics* (1992) 9:352. Panels for radiation hybrid mapping are available from Research Genetics, Inc., Huntsville, Alabama, USA. Databases for markers using various panels are available via the world wide web at <http://F/shgc-www.stanford.edu>, and other locations. The statistical program RHMAP can be used to construct a map based on the data from radiation  
5 hybridization with a measure of the relative likelihood of one order versus another, RHMAP is available via the world wide web at <http://www.sph.umich.edu/group/statgen/software>.

Such mapping can be useful in identifying the function of the target gene by its proximity to other genes with known function. Function can also be assigned to the target gene when particular syndromes or diseases map to the same chromosome.

#### 10 Tissue Profiling

The nucleic acids of the present invention can be used to determine the tissue type from which a given sample is derived. For example, a metastatic lesion is identified by its developmental organ or tissue source by identifying the expression of a particular marker of that organ or tissue. If a nucleic acid is expressed only in a specific tissue type, and a metastatic  
15 lesion is found to express that nucleic acid, then the developmental source of the lesion has been identified. Expression of a particular nucleic acid is assayed by detection of either the corresponding mRNA or the protein product. Immunological methods, such as antibody staining, are used to detect a particular protein product. Hybridization methods may be used to detect particular mRNA species, including but not limited to *in situ* hybridization and Northern  
20 blotting.

#### Use of Polymorphisms

A nucleic acid will be useful in forensics, genetic analysis, mapping, and diagnostic applications if the corresponding region of a gene is polymorphic in the human population. A particular polymorphic form of the nucleic acid may be used to either identify a sample as  
25 deriving from a suspect or rule out the possibility that the sample derives from the suspect. Any means for detecting a polymorphism in a gene are used, including but not limited to electrophoresis of protein polymorphic variants, differential sensitivity to restriction enzyme cleavage, and hybridization to an allele-specific probe.

#### B. Use of Nucleic Acids and Encoded Polypeptides to Raise Antibodies

30 Expression products of a nucleic acid, the corresponding mRNA or cDNA, or the corresponding complete gene are prepared and used for raising antibodies for experimental,

diagnostic, and therapeutic purposes. For nucleic acids to which a corresponding gene has not been assigned, this provides an additional method of identifying the corresponding gene. The nucleic acid or related cDNA is expressed as described above, and antibodies are prepared. These antibodies are specific to an epitope on the encoded polypeptide, and can precipitate or  
5 bind to the corresponding native protein in a cell or tissue preparation or in a cell-free extract of an *in vitro* expression system.

Immunogens for raising antibodies are prepared by mixing the polypeptides encoded by the nucleic acids of the present invention with adjuvants. Alternatively, polypeptides are made as fusion proteins to larger immunogenic proteins. Polypeptides are also covalently linked to other  
10 larger immunogenic proteins, such as keyhole limpet hemocyanin. Immunogens are typically administered intradermally, subcutaneously, or intramuscularly. Immunogens are administered to experimental animals such as rabbits, sheep, and mice, to generate antibodies. Optionally, the animal spleen cells are isolated and fused with myeloma cells to form hybridomas which secrete monoclonal antibodies. Such methods are well known in the art. According to another method  
15 known in the art, the nucleic acid is administered directly, such as by intramuscular injection, and expressed *in vivo*. The expressed protein generates a variety of protein-specific immune responses, including production of antibodies, comparable to administration of the protein.

Preparations of polyclonal and monoclonal antibodies specific for nucleic acid-encoded proteins and polypeptides are made using standard methods known in the art. The antibodies  
20 specifically bind to epitopes present in the polypeptides encoded by a nucleic acid of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-1103, even more preferably SEQ ID Nos. 1-503, and still more preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto. In a preferred embodiment the antibodies bind to  
25 epitopes on the polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493. Typically, at least about 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve noncontiguous amino acids may require more, for example, at least about 15, 25, or 50 amino acids. A short sequence of a nucleic acid may then be unsuitable for use as an epitope to raise antibodies for identifying the  
30 corresponding novel protein, because of the potential for cross-reactivity with a known protein. However, the antibodies may be useful for other purposes, particularly if they identify common structural features of a known protein and a novel polypeptide encoded by a nucleic acid of the invention.

Antibodies that specifically bind to human nucleic acid-encoded polypeptides should provide a detection signal at least about 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in Western blots or other immunochemical assays. Preferably, antibodies that specifically bind nucleic acid T-encoded polypeptides do not detect  
5 other proteins in immunochemical assays and can immunoprecipitate nucleic acid-encoded proteins from solution.

To test for the presence of serum antibodies to the nucleic acid-encoded polypeptide in a human population, human antibodies are purified by methods well known in the art. Preferably, the antibodies are affinity purified by passing antiserum over a column to which a nucleic acid-  
10 encoded protein, polypeptide, or fusion protein is bound. The bound antibodies can then be eluted from the column, for example using a buffer with a high salt concentration.

In addition to the antibodies discussed above, genetically engineered antibody derivatives are made, such as single chain antibodies.

Antibodies may be made by using standard protocols known in the art (See, for example,  
15 Antibodies: A Laboratory Manual ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, hamster, or rabbit can be immunized with an immunogenic form of the peptide (e.g., a mammalian polypeptide or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above).

In one aspect, this invention includes monoclonal antibodies that show a subject  
20 polypeptide is highly expressed in colorectal tissue or tumor tissue, especially colon cancer tissue or colon cancer-derived cell lines. Therefore, in one embodiment, this invention provides a diagnostic tool for the analysis of expression of a subject polypeptide in general, and in particular, as a diagnostic for colon cancer.

Techniques for conferring immunogenicity on a protein or peptide include conjugation to  
25 carriers or other techniques well known in the art. An immunogenic portion of a protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a protein  
30 of a mammal, e.g., antigenic determinants of a protein encoded by one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 or closely related homologs (e.g., at least 90% identical, and more preferably at least 95% identical).

Following immunization of an animal with an antigenic preparation of a polypeptide, antisera can be obtained and, if desired, polyclonal antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing  
5 cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256: 495-497), the human B cell hybridoma technique (Kozbar *et al.*, (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R.  
10 Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject polypeptides. Antibodies can be fragmented using  
15 conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)<sub>2</sub> fragments can be generated by treating antibody with pepsin. The resulting F(ab)<sub>2</sub> fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific, single-chain, and chimeric and humanized molecules having affinity for a polypeptide conferred  
20 by at least one CDR region of the antibody. In preferred embodiments, the antibodies, the antibody further comprises a label attached thereto and able to be detected, (e.g., the label can be a radioisotope, fluorescent compound, chemiluminescent compound, enzyme, or enzyme co-factor).

Antibodies can be used, e.g., to monitor protein levels in an individual for determining,  
25 e.g., whether a subject has a disease or condition, such as colon cancer, associated with an aberrant protein level, or allowing determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of polypeptides may be measured from cells in bodily fluid, such as in blood samples.

Another application of antibodies of the present invention is in the immunological  
30 screening of cDNA libraries constructed in expression vectors such as gt11, gt18-23, ZAP, and ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, gt11 will produce fusion proteins whose amino termini consist of  $\beta$ -galactosidase amino acid sequences and whose

carboxyl termini consist of a foreign polypeptide. Antigenic epitopes of a protein, e.g., other orthologs of a particular protein or other paralogs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with antibodies. Positive phage detected by this assay can then be isolated from the infected plate.

- 5 Thus, the presence of homologs can be detected and cloned from other animals, as can alternate isoforms (including splicing variants) from humans.

In another embodiment, a panel of monoclonal antibodies may be used, wherein each of the epitope's involved functions are represented by a monoclonal antibody. Loss or perturbation of binding of a monoclonal antibody in the panel would be indicative of a mutational attention of  
10 the protein and thus of the corresponding gene.

### C. Differential Expression

The present invention also provides a method to identify abnormal or diseased tissue in a human. For nucleic acids corresponding to profiles of protein families as described above, the choice of tissue may be dictated by the putative biological function. The expression of a gene  
15 corresponding to a specific nucleic acid is compared between a first tissue that is suspected of being diseased and a second, normal tissue of the human. The normal tissue is any tissue of the human, especially those that express the target gene including, but not limited to, brain, thymus, testis, heart, prostate, placenta, spleen, small intestine, skeletal muscle, pancreas, and the mucosal lining of the colon.

20 The tissue suspected of being abnormal or diseased can be derived from a different tissue type of the human, but preferably it is derived from the same tissue type; for example an intestinal polyp or other abnormal growth should be compared with normal intestinal tissue. A difference between the target gene, mRNA, or protein in the two tissues which are compared, for example in molecular weight, amino acid or nucleotide sequence, or relative abundance,  
25 indicates a change in the gene, or a gene which regulates it, in the tissue of the human that was suspected of being diseased.

The target genes in the two tissues are compared by any means known in the art. For example, the two genes are sequenced, and the sequence of the gene in the tissue suspected of being diseased is compared with the gene sequence in the normal tissue. The target genes, or  
30 portions thereof, in the two tissues are amplified, for example using nucleotide primers based on the nucleotide sequence shown in the Sequence Listing, using the polymerase chain reaction. The amplified genes or portions of genes are hybridized to nucleotide probes selected from a

corresponding nucleotide sequence shown SEQ ID No. 1-4494. A difference in the nucleotide sequence of the target gene in the tissue suspected of being diseased compared with the normal nucleotide sequence suggests a role of the nucleic acid-encoded proteins in the disease, and provides a lead for preparing a therapeutic agent. The nucleotide probes are labeled by a variety of methods, such as radiolabeling, biotinylation, or labeling with fluorescent or chemiluminescent tags, and detected by standard methods known in the art.

Alternatively, target mRNA in the two tissues is compared. PolyA<sup>+</sup>RNA is isolated from the two tissues as is known in the art. For example, one of skill in the art can readily determine differences in the size or amount of target mRNA transcripts between the two tissues using Northern blots and nucleotide probes selected from the nucleotide sequence shown in the Sequence Listing. Increased or decreased expression of a target mRNA in a tissue sample suspected of being diseased, compared with the expression of the same target mRNA in a normal tissue, suggests that the expressed protein has a role in the disease, and also provides a lead for preparing a therapeutic agent.

Any method for analyzing proteins is used to compare two nucleic acid-encoded proteins from matched samples. The sizes of the proteins in the two tissues are compared, for example, using antibodies of the present invention to detect nucleic acid-encoded proteins in Western blots of protein extracts from the two tissues. Other changes, such as expression levels and subcellular localization, can also be detected immunologically, using antibodies to the corresponding protein. A higher or lower level of nucleic acid-encoded protein expression in a tissue suspected of being diseased, compared with the same nucleic acid-encoded protein expression level in a normal tissue, is indicative that the expressed protein has a role in the disease, and provides another lead for preparing a therapeutic agent.

Similarly, comparison of gene sequences or of gene expression products, e.g., mRNA and protein, between a human tissue that is suspected of being diseased and a normal tissue of a human, are used to follow disease progression or remission in the human. Such comparisons of genes, mRNA, or protein are made as described above.

For example, increased or decreased expression of the target gene in the tissue suspected of being neoplastic can indicate the presence of neoplastic cells in the tissue. The degree of increased expression of the target gene in the neoplastic tissue relative to expression of the gene in normal tissue, or differences in the amount of increased expression of the target gene in the neoplastic tissue over time, is used to assess the progression of the neoplasia in that tissue or to monitor the response of the neoplastic tissue to a therapeutic protocol over time.



The expression pattern of any two cell types can be compared, such as low and high metastatic tumor cell lines, or cells from tissue which have and have not been exposed to a therapeutic agent. A genetic predisposition to disease in a human is detected by comparing an target gene, mRNA, or protein in a fetal tissue with a normal target gene, mRNA, or protein.

5 Fetal tissues that are used for this purpose include, but are not limited to, amniotic fluid, chorionic villi, blood, and the blastomere of an *in vitro*-fertilized embryo. The comparable normal target gene is obtained from any tissue. The mRNA or protein is obtained from a normal tissue of a human in which the target gene is expressed. Differences such as alterations in the nucleotide sequence or size of the fetal target gene or mRNA, or alterations in the molecular  
10 weight, amino acid sequence, or relative abundance of fetal target protein, can indicate a germline mutation in the target gene of the fetus, which indicates a genetic predisposition to disease.

In a preferred embodiment nucleic acid macroarrays comprising the one or more of the sequences of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488,  
15 4490, 4492, and 4494 may be used to evaluate differential expression of nucleic acid sequences in cancerous cells or tissue relative to the expression of the same sequences in normal cells or tissue as described above. Preferably, such sequences are differentially expressed by at least 3 fold in cancerous cells or tissue relative to normal cells or tissue. More specifically, the present invention provides the full length sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480,  
20 4482, 4484, 4486, 4488, 4490, 4492, and 4494 which are differentially expressed in cancerous colonic cells/tissue by at least 3 fold relative to normal patient samples. Thus, the sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, as well as the encoded polypeptides (SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, respectively) serve as valuable diagnostic markers for  
25 identifying and screening for colon cancer in a patient.

D. Use of Nucleic Acids, and Encoded Polypeptides to Screen for Peptide Analogs and Antagonists

Polypeptides encoded by the instant nucleic acids, e.g., SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, preferably SEQ ID Nos. 1-  
30 1103, even more preferably SEQ ID Nos. 1-503, and most preferably SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, or a sequence complementary thereto, and corresponding full length genes can be used to screen peptide libraries to identify binding partners, such as receptors, from among the encoded polypeptides. Preferably, the

polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493 may be used screen for binding partners.

A library of peptides may be synthesized following the methods disclosed in U.S. Pat. No. 5,010,175, and in PCT WO 91/17823. As described below in brief, one prepares a mixture of peptides, which is then screened to identify the peptides exhibiting the desired signal transduction and receptor binding activity. In the '175 method, a suitable peptide synthesis support (e.g., a resin) is coupled to a mixture of appropriately protected, activated amino acids. The concentration of each amino acid in the reaction mixture is balanced or adjusted in inverse proportion to its coupling reaction rate so that the product is an equimolar mixture of amino acids coupled to the starting resin. The bound amino acids are then deprotected, and reacted with another balanced amino acid mixture to form an equimolar mixture of all possible dipeptides. This process is repeated until a mixture of peptides of the desired length (e.g., hexamers) is formed. Note that one need not include all amino acids in each step: one may include only one or two amino acids in some steps (e.g., where it is known that a particular amino acid is essential in a given position), thus reducing the complexity of the mixture. After the synthesis of the peptide library is completed, the mixture of peptides is screened for binding to the selected polypeptide. The peptides are then tested for their ability to inhibit or enhance activity. Peptides exhibiting the desired activity are then isolated and sequenced.

The method described in WO 91/17823 is similar. However, instead of reacting the synthesis resin with a mixture of activated amino acids, the resin is divided into twenty equal portions (or into a number of portions corresponding to the number of different amino acids to be added in that step), and each amino acid is coupled individually to its portion of resin. The resin portions are then combined, mixed, and again divided into a number of equal portions for reaction with the second amino acid. In this manner, each reaction may be easily driven to completion. Additionally, one may maintain separate "subpools" by treating portions in parallel, rather than combining all resins at each step. This simplifies the process of determining which peptides are responsible for any observed receptor binding or signal transduction activity.

In such cases, the subpools containing, e.g., 1-2,000 candidates each are exposed to one or more polypeptides of the invention. Each subpool that produces a positive result is then resynthesized as a group of smaller subpools (sub-subpools) containing, e.g., 20-100 candidates, and reassayed. Positive sub-subpools may be resynthesized as individual compounds, and assayed finally to determine the peptides that exhibit a high binding constant. These peptides can be tested for their ability to inhibit or enhance the native activity. The methods described in WO

91/7823 and U.S. Patent No. 5,194,392 (herein incorporated by reference) enable the preparation of such pools and subpools by automated techniques in parallel, such that all synthesis and resynthesis may be performed in a matter of days.

Peptide agonists or antagonists are screened using any available method, such as signal  
5 transduction, antibody binding, receptor binding, mitogenic assays, chemotaxis assays, etc. The methods described herein are presently preferred. The assay conditions ideally should resemble the conditions under which the native activity is exhibited *in vivo*, that is, under physiologic pH, temperature, and ionic strength. Suitable agonists or antagonists will exhibit strong inhibition or  
10 enhancement of the native activity at concentrations that do not cause toxic side effects in the subject. Agonists or antagonists that compete for binding to the native polypeptide may require concentrations equal to or greater than the native concentration, while inhibitors capable of binding irreversibly to the polypeptide may be added in concentrations on the order of the native concentration.

The end results of such screening and experimentation will be at least one novel  
15 polypeptide binding partner, such as a receptor, encoded by a nucleic acid of the invention, and at least one peptide agonist or antagonist of the novel binding partner. Such agonists and antagonists can be used to modulate, enhance, or inhibit receptor function in cells to which the receptor is native, or in cells that possess the receptor as a result of genetic engineering. Further, if the novel receptor shares biologically important characteristics with a known receptor,  
20 information about agonist/antagonist binding may help in developing improved agonists/antagonists of the known receptor.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such  
25 techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press:1989); *DNA Cloning*, Volumes I and II (D.N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al.* U.S. Patent No. 4,683,195; *Nucleic Acid Hybridization* (B.D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods in Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M.P. Calos

eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C.C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

As mentioned above, the sequences described herein are believed to have particular utility in regards to colon cancer. However, they may also be useful with other types of cancers and other disease states.

The present invention will now be illustrated by reference to the following examples which set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and are not to be construed as restricting the invention in any way.

## XI. Examples

### A. Identification of differentially expressed sequences.

#### Description of the Libraries

SEQ ID Nos: 1-4470 were derived from libraries designated as 101, 102, 103, 104, 109, 110, 111, and 112 as described below briefly and in the accompanying table (Table 1). For example, the 101 library is a normalized, colon cancer specific, subtracted cDNA library. It is specific for sequences expressed in colon cancer [proximal and distal Dukes' B, microsatellite instability negative (MSI-)] but not expressed in normal tissues, including normal colon tissue. The 102 library is a normalized, colon specific, subtracted cDNA library. It is specific for sequences expressed in normal colon tissue but not expressed in other normal tissues. Characteristics of the remaining libraries are described in Table 1.

**Table 1 Library designation and description**

Library Designation	Description
101	Specific for sequences expressed in colon cancer (proximal and distal Dukes' B, MSI-) but not expressed in normal tissues <sup>4</sup> , including colon <sup>2</sup>
102	Specific for sequences expressed in normal colon (normal tissue from proximal and distal Dukes' B, MSI-matrix patients) <sup>3</sup> , but not expressed in

	other normal tissues <sup>4</sup>
103	Specific for sequences expressed in proximal Dukes' B, MSI- colon cancer, but not expressed in normal colon tissue <sup>3</sup>
104	Specific for sequences expressed in distal Dukes' B, MSI- colon cancer, but not expressed in normal colon tissue <sup>3</sup>
109	Specific for sequences expressed in proximal Dukes' B, MSI+ colon cancer, but not expressed in normal colon tissue <sup>3</sup>
110	Specific for sequences expressed in proximal Dukes' B, MSI+ colon cancer, but not expressed in other normal tissues <sup>4</sup> , including colon <sup>2</sup>
111	Specific for sequences expressed in distal, Dukes' D, MSI- colon cancer, but not expressed in normal colon tissue <sup>3</sup>
112	Specific for sequences expressed in distal, Dukes' D, MSI- colon cancer, but not expressed in normal tissues <sup>4</sup> , including colon <sup>2</sup>

<sup>1</sup> cDNA synthesized from SW480 poly A+ RNA obtained from Clontech, Palo Alto, CA

<sup>2</sup> cDNA synthesized from normal colon tissue total RNA obtained from OriGene Technologies, Inc.; Rockville, MD

<sup>3</sup> Corresponding normal colon epithelium from colon cancer patients.

- 5 <sup>4</sup> A pool of cDNAs synthesized from the following normal tissue RNAs (poly A+ or total) obtained from OriGene Technologies, Inc.: heart, kidney, spleen, liver, peripheral blood lymphocytes, small intestine, skeletal muscle, lung and prostate.

#### Construction of the normalized and subtracted cDNA libraries

10 The normalized and subtracted cDNA libraries were constructed according to published procedures (Daitchenko et al., 1996 PNAS 93:6025-6030, Gurskaya et al., 1996 Analytical Biochemistry 240:90-97). Commercially available kits from Clontech Laboratories, Inc., Palo Alto, California were utilized (Clontech SMART cDNA synthesis kit, catalog number K1052-1, and Clontech PCR-Select cDNA Subtraction kit, catalog number K1804-1). For each subtracted library, the specific or "tester" cDNA was comprised of amplified cDNA from four similar  
15 sample types that were pooled together. Likewise, the reference or "driver" cDNA was comprised of a pool of sample types as illustrated in Table 1. During the subtraction process, the genes or transcripts unique to the tester are retained, and the genes or transcripts common to both the tester and driver are removed. Thus, in principle, the clones present in the subtracted libraries indicate those genes or transcripts that are expressed (or overexpressed) in the tester, but

not expressed (or underexpressed) in the driver. Reverse-subtracted libraries were also constructed in which the tester and driver materials were reversed. These libraries were only utilized to prepare labeled targets (see below).

To construct the libraries, one microgram of total RNA from each sample was  
5 representatively amplified using the Clontech SMART cDNA synthesis kit. The amplified cDNA was purified and pooled to create the individual tester and driver samples that were used for the subsequent library construction. To construct the normalized and subtracted libraries, the Clontech PCR-Select cDNA Subtraction kit was utilized. A forty-five fold mass excess of driver cDNA (450 nanograms) was used for each subtraction experiment. Subtractive hybridization of  
10 tester with driver cDNAs was performed twice, each time for about 8-12 hours. Subtracted cancer specific cDNA was ligated into the pCR2.1-TOPO plasmid vector (Invitrogen Corporation, Carlsbad CA) and chemically transformed into ultracompetent Epicurian E. coli XL10-Gold cells (Stratagene, La Jolla, CA). The transformed cells were plated onto LB-ampicillin plates containing IPTG and X-gal. Individual white colonies, representing those with  
15 cloned inserts, were picked and grown overnight in LB-ampicillin broth. Plasmid DNA was purified using QIAprep 96 Turbo kits from Qiagen (Valencia, CA).

#### Sequencing of the clones

The nucleotide sequence of the inserts from clones was determined by single-pass sequencing from either the T7 or M13 promoter sites using fluorescently labeled  
20 dideoxynucleotides via the Sanger sequencing method. The nucleotide sequences of the individual clones were compared to those in public databases (GenBank, dbEST, Geneseq) via Blast 2 homology searches according to methods described in the text.

The sequences derived from individual clones from the libraries described above represents a sequence from a partial mRNA transcript, since the cDNA used for making the  
25 subtracted library was restricted with *Rsa*I, a four base cutter restriction endonuclease that generates fragments with an average size of about 600 base pairs.

The nucleic acids of the invention were assigned a sequence identification number (see Figure 1). The nucleic acid sequences are provided in the attached Sequence Listing.

#### Validation of differential expression in colon cancer

30 To validate that the differentially expressed sequences found in this library were specific to colon cancer, the inserts from the plasmid DNA were amplified by PCR using vector-specific

primers. The amplification products were arrayed onto nylon membranes and hybridized with <sup>33</sup>P-labeled cDNAs prepared from both the subtracted library cDNA as well as the corresponding reverse-subtracted cDNA library. Each membrane array comprises approximately 3,456 clones. Four such membranes were generated comprising the clone libraries shown in Table 1 as indicated below in Table 3.

Membrane ID Number	Library Clones
101-1	Clones from subtracted library 101
101-2	Clones from subtracted library 101 and 102
103104109	Clones from subtracted libraries 103, 104, and 109
110111112	Clones from subtracted libraries 110, 111, and 112

The set of four membranes is hybridized, using techniques known to those of skill in the art and further described above, with <sup>32</sup>P-labeled target nucleic acid molecules obtained from human colon cancer tissue. A second, identical set of membranes is hybridized with <sup>32</sup>P-labeled target nucleic acid molecules obtained from normal human colon tissue. The signals of the hybridization products on the cancer membrane are subsequently compared to those on the normal membrane. A difference in hybridization, indicative of a difference in expression of the sequence in colon cancer vs. normal, of at least 3 fold is considered to be indicative of differential expression.

Using this validation technique, the full length cDNA sequences of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 have been identified as significantly differentially expressed in colon cancer relative to normal colon tissue.

Those skilled in the art will recognize, or be able to ascertain, using not more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such specific embodiments and equivalents are intended to be encompassed by the following claims.

All patents, published patent applications, and publications cited herein are incorporated by reference as if set forth fully herein.

What is claimed is:

CLAIMS

1. A method for detecting cancer in which one or more of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 are used as probes,  
5 said method comprising:
  - (a) collecting a sample of cells from a patient,
  - (b) isolating nucleic acid from the cells of the sample,
  - (c) contacting the nucleic acid sample with one or more primers which specifically hybridize to a nucleic acid sequence of SEQ ID Nos. 1-4470, 4472, 4474, 4476,  
10 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494 under conditions such that hybridization and amplification of the nucleic acid occurs, and
  - (d) comparing the presence, absence, or size of an amplification product to the amplification product of a normal cell.
2. A method of claim 1 in which said cancer is colon cancer.
- 15 3. A method for detecting cancer in a patient sample in which an antibody to a protein encoded by SEQ ID Nos. 1-4470 is used to react with proteins in said sample.
4. A method for detecting cancer in a patient sample in which an antibody to a protein encoded by one or more of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, or 4494 is used to react with proteins in said sample.
- 20 5. A method for detecting cancer in a patient sample in which an antibody to a protein having the sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, or 4493 is used to react with proteins in said sample.
6. A method for identifying an agent which alters the level of expression in a cell of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470 or a  
25 sequence complementary thereto, comprising
  - (a) providing a cell;
  - (b) treating the cell with a test agent;



(c) determining the level of expression in the cell of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470 or a sequence complementary thereto; and

5 (d) comparing the level of expression of the nucleic acid in the treated cell with the level of expression of the nucleic acid in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of expression of the nucleic acid in the untreated cell is indicative of an agent which alters the level of expression of the nucleic acid in a cell.

7. A method for identifying an agent which alters the level of expression in a cell of  
10 a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, or 4494 or a sequence complementary thereto, comprising

(a) providing a cell;

(b) treating the cell with a test agent;

15 (c) determining the level of expression in the cell of a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, or 4494 or a sequence complementary thereto; and

(d) comparing the level of expression of the nucleic acid in the treated cell  
20 with the level of expression of the nucleic acid in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of expression of the nucleic acid in the untreated cell is indicative of an agent which alters the level of expression of the nucleic acid in a cell.

8. A method for identifying an agent which alters the level of expression in a cell of  
a polypeptide of one or more of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485,  
25 4487, 4489, 4491, or 4493 comprising

(a) providing a cell;

(b) treating the cell with a test agent;

(c) determining the level of expression of one or more polypeptides of SEQ  
ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, or 4493 in said cell

by reacting said cell with an antibody specific for one or more of the polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, or 4493; and

- (d) comparing the level of expression of said one or more polypeptides in the treated cell with the level of expression of said one or more polypeptides in an untreated cell, wherein a change in the level of expression of the nucleic acid in the treated cell relative to the level of expression of the nucleic acid in the untreated cell is indicative of an agent which alters the level of expression of the polypeptide in a cell.

9. A pharmaceutical composition comprising an agent identified by the method of claim 29, 30, or 31.

10. A pharmaceutical composition comprising a nucleic acid which includes a nucleotide sequence which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470 or a sequence complementary thereto.

11. A pharmaceutical composition comprising a polypeptide encoded by a nucleic acid which includes a nucleotide sequence that hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470 or a sequence complementary thereto.

12. A pharmaceutical composition comprising a polypeptide having the sequence of one of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, or 4493.

13. A pharmaceutical composition comprising an antibody which binds to one or more polypeptides having the sequence of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, or 4493.

14. A method of determining the phenotype of a cell, comprising detecting the differential expression, relative to a normal cell, of at least one nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, wherein the nucleic acid is differentially expressed by at least a factor of two.

15. A method for determining the phenotype of cells in a sample of cells from a patient, comprising:

(a) providing a nucleic acid probe comprising a nucleotide sequence having at least 12 consecutive nucleotides of any of SEQ ID Nos. 1-4470, 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494;

(b) obtaining a sample of cells from a patient;

5 (c) providing a second sample of cells substantially all of which are non-cancerous;

(d) contacting the nucleic acid probe under stringent conditions with mRNA of each of said first and second cell samples; and comparing (a) the amount of hybridization of the probe with mRNA of the first cell sample, with (b) the amount of hybridization of the probe  
10 with mRNA of the second cell sample, wherein a difference of at least a factor of two in the amount of hybridization with the mRNA of the first cell sample as compared to the amount of hybridization with the mRNA of the second cell sample is indicative of the phenotype of cells in the first cell sample.

16. A method of determining the phenotype of cell, comprising detecting the  
15 differential expression, relative to a normal cell, of at least one polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-4470, wherein the polypeptide is differentially expressed by at least a factor of two.

17. A method of determining the phenotype of cell, comprising detecting the  
20 differential expression, relative to a normal cell, of at least one polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to a sequence selected from the group consisting of SEQ ID Nos. 4472, 4474, 4476, 4478, 4480, 4482, 4484, 4486, 4488, 4490, 4492, and 4494, wherein the polypeptide is differentially expressed by at least a factor of two.

18. A method of determining the phenotype of cell, comprising detecting the  
25 differential expression, relative to a normal cell, of at least one polypeptide selected from the group of polypeptides of SEQ ID Nos. 4471, 4473, 4475, 4477, 4479, 4481, 4483, 4485, 4487, 4489, 4491, and 4493, wherein the polypeptide is differentially expressed by at least a factor of two.

19. The method of claim 16, 17, or 18, wherein the level of said polypeptide is detected in an immunoassay.

20. A method for detecting a mutation in a test nucleic acid which hybridizes under stringent conditions to a nucleic acid of SEQ ID Nos. 1-4470 or a sequence complementary thereto, comprising

- (a) collecting a sample of cells from a patient,
- 5 (b) isolating nucleic acid from the cells of the sample,
- (c) contacting the nucleic acid sample with one or more primers which specifically hybridize to a nucleic acid sequence of SEQ ID Nos. 1-4470 under conditions such that hybridization and amplification of the nucleic acid occurs, and
- 10 (d) comparing the presence, absence, or size of an amplification product to the amplification product of a normal cell.

21. An isolated nucleic acid comprising a portion of a nucleotide sequence of SEQ ID Nos. 504-1103 or a sequence complementary thereto.

22. A gene which hybridizes to one of SEQ ID Nos. 1-503.

15 23. An isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 1-503 or a sequence complementary thereto.

24. An isolated nucleic acid comprising a nucleotide sequence at least 80% identical to a sequence corresponding to at least about 15 consecutive nucleotides of one of SEQ ID Nos. 1-503 or a sequence complementary thereto.

20 25. An isolated nucleic acid comprising a nucleotide sequence of SEQ ID Nos. 1-503 or a sequence complementary thereto.

26. A nucleic acid according to claim 25, further comprising a transcriptional regulatory sequence operably linked to said nucleotide sequence so as to render said nucleotide sequence suitable for use as an expression vector.

25 27. An expression vector, capable of replicating in at least one of a prokaryotic cell and eukaryotic cell, comprising the nucleic acid of claim 26.

28. A host cell transfected with the expression vector of claim 27.

29. A transgenic animal having a transgene of the nucleic acid of claim 25 incorporated in cells thereof, which transgene modifies the level of expression of the nucleic acid, the stability of an mRNA transcript of the nucleic acid, or the activity of the encoded product of the nucleic acid.,

5 30. A substantially pure nucleic acid which hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 12 consecutive nucleotides of one of SEQ ID Nos. 1-1103 or a sequence complementary thereto.

31. A polypeptide including an amino acid sequence encoded by a nucleic acid of claim 25 or a fragment comprising at least 25 amino acids thereof.

10 32. A probe/primer comprising a substantially purified oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least 12 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos. 1-1103.

15 33. An array including at least 10 different probes of claim 32 attached to a solid support.

34. The probe/primer of claim 32, further comprising a label group attached thereto and able to be detected.

35. The probe/primer of claim 34, wherein said label group being selected from radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors.

20 36. An antibody immunoreactive with a polypeptide of claim 31.

37. A method for determining the presence or absence of a nucleic acid which hybridizes under stringent conditions to one of SEQ ED Nos. 1-1103 in a cell, comprising contacting the cell with a probe of claim 32.

25 38. A method for determining the presence of absence of a polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to one of SEQ ID Nos. 1-503 in a cell, comprising contacting the cell with an antibody of claim 36.

39. An antisense oligonucleotide analog which hybridizes under stringent conditions to at least 12 consecutive nucleotides of one of SEQ ID Nos. 1-503 or a sequence complementary thereto, and which ~ resistant to cleavage by a nuclease.

40. A test kit for determining the phenotype of transformed cells, comprising the probe/primer of claim 34, for measuring a level of a nucleic acid which hybridizes under stringent conditions to a nucleic acid of SEQ ID Nos. 1-4470 in a sample of cells isolated from a patient.

- 5           41. A test kit for determining the phenotype of transformed cells, comprising an antibody specific for a protein encoded by a nucleic acid which hybridizes under stringent conditions to any one of SEQ Nos. 1-4470.

Figure 1

SEQ ID NO: 1      GGTACATTGAATTACAAAAGGATCCAAGAATATTGAAATAGTTACCAAAAAA  
 ATTTGATATAGAAATATATATGGTTTATTAATGAATAAGATCTAGCAGTGGCCTTACTATAATTTT  
 AACACAGTAGCGTGGTGAGCATAAATATTTCAACTTATCTGCAACAATTGTAATGTGATATGGAA  
 ATAATCAGTTTAATTGTGACAGTCACATTTACTGCTAACTACTGTGATTGATTGGCTATGTTTACAT  
 TTTATTAGAAGTTATTGAAAGTAGATGCCAGGCATGGTGGCTCATGCCTGTAATCCTAGCACTTTG  
 GGAGGCCAAAGTGGGCAGATCAITTTGAGCCACGAGTTGAGACCAGCCTGGCCAACATGATNAA  
 GCCCTGTCTCTACAANANAAANTACNANTTNNGNANANAAAATGTTNCCTGCCCGGNCCGGCCGTT  
 CNAAAGGGCGAAATTC AACNACACTTGGCGGGGCGTTTNTAANTGGTNCGAACCTTNGGTNCCN  
 AGCTTTGGCGTNAATCATGGNTNATN

SEQ ID NO: 2      GGTACTCAGTATCTTGACTGACTTGTGGAAATACACTTTAGATTATTTACGT  
 GGAATATATTCAGTTAATTGAGCAAAAGTATATGTATTGCTGTGGAGATTATTTTGTCTTTTCTTG  
 ATTATATTTATAACTTAGTGAAAAGGTAGAGAAAAGTATTACTGAAAAAATTACATATTTTATAGGT  
 GACTTGTATTTGACTTTTGCATTTCTAATTTAATTTCTTGTAGATGGCTGACCACATTTCAAAAA  
 TTGCCATCCTCTTTATCTCTGTTTCTTATACAAACATTTTCTTTCAAAATTGTTTCTAGAGGTTTGGG  
 AAAAGAAGCATAAGTAATTAAGGAGCAAAATTCTGGGCCTAATATTATTGCTCCTTTATAAGCT  
 ATAATCTTTTTATTTAAATTTTTNTGNGGCCCTGGATTTTAACTTTGANTCATTTNTTAGNGGCTT  
 TANNNTATCATTNCCCCNGGCTNCNTNAAGGTTNCCCNAGATNAAAAGGTNNNNNTGNNNNNTNN  
 NCANAANCTNTTAANATTCCTTNNNGGNNNNNTTNNNNNNNNNTNNNNNANGGTTNAAATNCC  
 TNNCTNNNC

SEQ ID NO: 3      GGTACAATTGTCTTTTCTGGGACACTCACTTCTGAAAAGAAGGCAGGAATTG  
 GAAGGGCTGAAAAAGGCATCGTGATGAAATCCACGTCCTGCCAAGTTGTACTGTAAAGTTTTAGT  
 CCCGGCAGTCAAGGCCACAATAGACAAAAATAAAGTCAGGAGGGAAGCCCACTGGATCCAGTTT  
 AGACGCCTCTTCAGCACTATCCTGAATAGAAGAGCTGTTGTTATAATGCTAAAATTTGAGAAGATA  
 ACAGCCATGGCTGGTTGAAGATAGGACAGGACATAGAAGACAATCAAGTTATCCAGGAAATAAA  
 GAAAGGCAGGAATGGACCACTTCATGAAATCAGAGAATTCCTTCCAGGAAGCATATTTCAAATTT  
 CTACTTTGATGATCTTTCTTTATTACNCAGAATGGCCCCAGCCCCCNGNAACTAACTTTAACCAG  
 TTTTGACCCCCCNTTNCAGTAGTTGGAANAAAACAAACCTGGTTTCTTCATTGGCGAATCCTTCCTT  
 N

SEQ ID NO: 4      ACTTNGANACACAATTTTAC  
 TNTGTNCCCCAGGCTGGAGTGCAGNGNGTGATCTCAGCTCACTGCAACTTNTGCCTCCTGGGTTC  
 AAGCAATTNTCATGCCCCANCCTCCCAAGTAGCTGGGAATACAGTTGNGTGCCNCCCACTGGCT  
 TATTTTTTGNATTTTTTTANNAANATGGGGTTTCAACCATGTTGGCCAGGCTGGNCTNTAACTNTTA  
 TCCTCAAGNGATCTGCCTGCCTTGGCCTCCCAAAGNGCTGGGATTACAGGCATGAGCCACCGCAC  
 CTGGCCTACTTATCCTGTAAATGAAAATATTTGATTGGACAAAACATCTTATTCCTTAACTCTGNTA  
 TGNGCTTTTGAAAATGGTTTTACCTACTGGANAANTTTGNTGGTCCCAAANGCCCCAAAATNNNT  
 TGAANAGTTCCCCCNTTCTNAAANGGGANAAAAAAATTTTNTTTAATAGGGGN

SEQ ID NO: 5      ACGCGGGCCCGGTGATGCCATCCTCACAGTGTTAAGTGACTTGGAAGATGAG  
 ACCATGAAAAGAACAAGTGCTGGTGGCACTCCCAGGCAAAGATCCCCAGGCTGGAGGAGTTGGC  
 GCCTGAAGAAAACCAGAGCAAAACCTCAAGGGTCAGAGGGCCTGGGCATCGGTGCAGGGCTCAC  
 TTGAGACTGAACACGTTCCCCAGGGAAGATCTGTATGCTTCTAAAGAACACTTTTGCCAGGCACA  
 GTGAATCATGCCTGTAATCCCAGCACTTTGAGAGGCTGANGCAGGTGGAATGCTTGAGCTCAGGA  
 GTTCGAGACCAGCCTGGGCAACATAGCAAAACCTTGTCTCTACTAAAAATCAAAAAAATTAGCCC  
 AGATGTGGTGGGTGCCTGGCTGTAGTCCAGCTACTTANGAACTGANGTAGGAAGATCACTTTG  
 AGTCTGGGAAAATGAAGCTTCAATGAACCTTGANTACACCCGNTTGACTTCANCTGAGTGACNG  
 AAGGAGACCTTGTTTGAAAAAAAANNNAAAAAAAGTCCCTTGCCCGGAACACCCCTTANGG  
 NGAATTCNCCCACTGGGGGCCGTCTAANGGANCCACTTGGACCAACTTGGGNAAAATGGGATAA  
 TTN

SEQ ID NO: 6      GGTACTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTGACTTTTGATTTTC  
 TCACTGTGTTCTTTTGGTATTGTAAAAATAGTAAATGTTAAAAAAGGAAAAAGAA  
 AAAGAAAATNTACTTTACAAAAAGTAANTTACAAAGGACTACTACTNATCTTTATGTTCAAATTTA  
 TAACAAAATGGGAAACNANTTAGGATGCTTCCATAGCCAGGAAAGGCAATGGTNANAATTTTC  
 CAGCTGAGGTTTGTGATCTGCTATTTTCCACATTAAGTNAAAAAANTCAAAACAATGTCTCAACA  
 NAANCAANAAAAATANTCCNTCATTCTTTTTGATTATATNNANAAGNGATTAACCNAAGTATGTT  
 ANTTTANTAAATAATNAAAAACATTTTNAANAGNNGGAACCNCTTCTAATTAACCATATCTNGG







ATCAATCAAAAATTGNGAATGCTAA

SEQ ID NO: 22 AATTCGCCCTTAGCGTGGTCGCGGCCGAGGTACTTTTTTTTTTTTTTTTTTTTTT  
TTNTTTTTNAAANAGGGGAAACATGTTTATTTTTATTNTTTAGGGTTCCAAANACTGGGGTAGGTA  
AAACTNTTGAANATTAACAAGGCAAACCTCACCANANAAAANAGTGTCCNGNTTAGANTTGANTTG  
GAAAGATGGNCCGGCTTTATTTTTGACNCCCCTTNGTATTGCTGNATCANNTCTTATACATGTATTT  
NTNCAACTTGCCNTCTTTTTTTTNA

SEQ ID NO: 23 ANTCGCCCTTAGCGTGGCCNCGGGCGAAGTACAAATCACTTGNTAGGCCTC  
AGTTTCTGCNACTNTGAANATCACTAGATTGCTAGCTNGTCTCTAAAATNGCTGCGACTCCANA  
TACTTNGCACTGAACAGAATCTAGGGTGTATGATATCTGTTTCANCTAAAGGCTNACGGGAAATGT  
CTA

SEQ ID NO: 24 GGTACGCGGGACACCAGCTCCGAAATCACCACCAAGGACTTAAAGGAGAAG  
AAGGAAGTTGTGGAAGANGCANAAAATGGAANAGACGCCCTGCTAACGGGAATGCTAATGACG  
AAAATGGGGAGNCACGATGCNTGACNATGAGNGTTTACCAA

SEQ ID NO: 25 CACCCTTACCGTGGTCNCNGCCNAGGTACTGGATGAAGCTGACCAAATGTTA  
ANCCGNGGATTCAAGGACCANATCTATGACNTATTCCAAAAGCTCAACAGNAACACCCANGTAGT  
TTTGCTGTCNCCCACACNNCCTTCTGATGTGCTGGAGGGGNNCAAGAATNAGCTGACTGGACCCC  
ATTGCGATTCTTCTCACTGAACCGANTTTAGACCCTNAACGGTGTCTGCTGCCAGCTGTACTTAATN  
AATCGAACCAACGGCGACGTGGAAGAGGNGGGCACNCTNTATTGANGCTNTCTTANACTTTNTCC  
ATCACTNGACTCAGTNNTCTTCGGTCAACNCCCCTGAGGAAGAGCGGNCATGCACTCTCACTCAA  
AGAGCNGTNGCTNCCGGGATTTCACNGTATTCA

SEQ ID NO: 26 ACGGAATAGAATGGAATGGAACGAATTGTAATGGAATGGAATGGAATGCAA  
TGGAATGGAATGGAATCAACGTGAGTGCAGGGGAATGTAATGGAACGGAATGCAATGGAATGGA  
ATCATCCGATTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATCAACCCG  
AGTGCAATGGAATGGAGTGAATGGAATGGAACCAACANAATGGAATGAAATGTAATGGAGAGT  
AAGGGAGTTGAATAGAATCAATCGGAATGTAATGGAATGGAATGCAACGGAATGGAATGGAATG  
GAATGGAATGGAATGGAATGGAATGGAATGGAATGATACGGAATAGAATGGAATGGAACGAAAT  
GGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATCGTTCCGAGTGAATAGGAGGGAA  
TGTATTCCANTGNAATTGGAAGGAATGGAATCAACCCANAGTGAATGGAAGGGAATGGGAAT  
GGAATGGAACCTAATAGAATANAATCNCNACAGGAATTTAATTGGAANGACTGGATGGAAA  
TGGAATGGGAATGGAATCAAACTCCATTGGAATTGGAATGGATTCCNCNTCCCTNNGGGCCCT

SEQ ID NO: 27 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTCGCGANATGGAGTCTTGCTCTGTTGCC  
CAGGCTGGAGTGCANCAAAGTGATCTCAGCTNACCGCAACCTNTGCCTCCCGGGTTCAAGCAATT  
NTINTGCCTCANCNTNTTGAAGTGGGACTACAGGCGCNCNCCCCCATGCCTGGCTAATTTTTGT  
ATTTTTAGTAAANAAGGAGTTTCACCATNTTACGAGGATGATCTCNATCTCCTGACCTTGNATC  
CACCTGCCTCCCCGGNCAATGGCATGATCACTGCTTACTGCAACCTGAAACCTCCTGGGCTAAGCG  
ATCCTTTCNCGANTGGCTNANACTACAGGAGCNCACCCANCNGGCCAGTTAATTTTTAAATTATT  
TATAA

SEQ ID NO: 28 AATTCGCCCTTAGCGTGGTCGCGGCCGAGGTACTGGAGGATTTATATATCA  
GCTATTGACTAATTCCAATCAGTGAATTAGCTTAGCCTAGGTAGATCCAATTAGCCAGTTATCTAA  
ATAGCAACTGCTGATTCATCAAGGCTGTTCCAAAATAGTGGATGTGAGCAGCACATGCNTACACC  
AAGAAGGAGAGCAGGGCCCTGAAATACTAACCTGGACAGAAAGCATCACGTGGCTTTACCATCAG  
CTTAAACTCTCCACTATCAGGGTTTAGTTTTGCCTGCTTCTAAGNCCTAAACTGGCANNNAACCTTA  
ACTCAGATCANCNTTACTTTTTANAGGTNNNGAGTGGAGCANNNGGGTAAATTCGAANATGAACC  
TA

SEQ ID NO: 29 CNGGGCGATTGGGCCCTCTATATGCATGCTCGAGCGGGCGCCAGTGTGATGG  
ATATCTGCAGAATTCGCCCTTGTGGTTCGCGGCCGAGGTTCTTTGGCCTCTNTGGGATANAAGTTAT  
TCAGNAGGCACACAAACAGAGGCNTTINCANANTTNAACTGCTCATNATATGGNNGNAAGATGA  
AGACANGATGGTNGGCCACNGTNCGNATTNATTTCCACCTTGGTNCCTTGNACGANCNTTTCNTNG  
GTTTGTATAANCNTNATCATCGCTTTAATAANCCCT

SEQ ID NO: 30 NGCATGCTCGACGGCCGCCAGGGNGANGGGTGATGGATATCTGCAGAATTTCG  
CCCTTAGCGTGGTTCGCGGCCGAGGTACTTCATTATAAGTAAGGTGTCTCTAAAAGGGACAGATCTC  
CTAGACCCCTCCTTAACCAAGTANCCAGTCTCTGATATCATTAATGGGTGATGGACAAACTAATACC

TTCTCTGCCCCGAGATGGGCTGAGGNIGGAACTCNCACCATTGTCTTCTTGCCANGTGGTCCCCGG  
CCAAACGTTTAAGGCTGGATTTTTTAATCCATNGGAAGACATTTTTCAGACAACTTNCANAAATTGT

SEQ ID NO: 31 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGANACGGAGTCACCTAGGCTGGCATGC  
AGGGATTTGATACTGGCTCACTGTAACTCAGCCTCCTAGGCTCAAGCGATTCTCTGCTTCAGCC  
TCTTGAATAGCTGGGATTTTCAAGGCATCTGCCACCACTCGTGGCTAAATTTTGGATTTTCAGTANAG  
ACCGGATTTCTCCATCATGGCCAGGCTGGTCTCAAACCTCTGATCTCTGGNGATCTGCCCACTTCG  
GCCTCCCAAAGTGCTGGGATTACAGGCATGAGTCACCGCCCCCGGCCTCATTGAAAAATTTATTTT  
TNAATACCNAACTGGATTGGTTNTTTTGGGGCAAAAANCCTTTTNNCNAATGNGCTTTTTTTTNC  
AAGTATTTTTTTTNAATTTCANANAANNTTTAAATTTTTTNGATANATTNGAAAAAANTANATTTGGT  
TAGTTNNAATA

SEQ ID NO: 32 ACGCGGGACTCAAAGAGACTAACAGTATTGTAAATCTAAGCTCTGTAAAGA  
AATTCCAAGTTAGTTTAACTACAGAGCTACAAAAATGTCACAGAAAATTGTTCCCTAGTGGCAATAT  
TAAAGAAAATAAAAAATTATTAGGCCAGGCACGGTGGCTCACACCTGTAATCCTTGCATTTGGGAG  
GCTGAGGTGGGCGGATCACCTGTGGTCAGGAGTTCAAACCCAGTCTCTACTAAAATACAAAAATT  
ANCCAGGTGTGGTGGCCATGCCTGTAATCCAGCTACTTGAGGGGCTGAAGTATGGANAATCNNTT  
TGAACCCANGANGTGGAGGCTNGNANTAAGCCCAAAATGNGGCCNNTTGTATT

SEQ ID NO: 33 AATTGNGCCCTCTAGATGCCANGCTCGAGNCGGGGCGAATTCGAGCTCGGTA  
CCCGGGGATCCTCTANAGTCNACCTGCAGGCATGCAAGCTNACNTNTTCNATTGAGAAGCCCCAA  
CAGCNTTGNNGGNCATCATNGACCTGGCTNCCCTCCTGCACTGAAAAA

SEQ ID NO: 34 ACGCGGGGGCNGTGCTGTTGGGAGTTGCTTGGAGGTNGGCGGCGCNGGGCTN  
AANGCTAGCAAACCGAGCGATCATGTNGCACAACAAATTTACTATTTCGGACAAATACAACTACN  
AGGAGTTTGNGTATCNACATGTNATGCTGCCNAGGACATATCCA

SEQ ID NO: 35 ACAGCACTCCATTACACAGAGTAACCCCACTCTTGATTAATCTGTTCTAAAG  
TGCCAGTATTATTTACACTTTTTTTTTTTTTTANCCAAAAGTCTGGCCAGTTGTGGCATCAGGTGAA  
GATTGTCAATCCANCTCTATTATCATTNACATTNANCAAGGGAAATTCTNATAATNCAGNTCTATG  
TCCCTGGTCCCGAAGGTTTACNTNGNCATTGGCANCNCTAAANTGGNGAACTTNTTCCNGNTCTN  
GGANCNTGAAAGNGGAGNCNAGGTANTGGCTGTTCAAAGG

SEQ ID NO: 36 ACCATTNTATTTAGTGTTGTAGGAAATGTTGGGTTACTTCTTAAAAACGAAAC  
CAAAGAAATTCAAAAGTCCCAAAGAAAGAAAACAGGAAATAATAATTCTATAATCCAAAAACGT  
TGGGCGATCCTTCAATNGGAGGAANANGGCGTCANTTAANTAGCTCACACTGTANATNTGGANAC  
ACCATATGGANATACGGAGTTAAGNTNGGGTGGATACTAGGAATTAANTTCTCCCCCTAANGC  
TAAATNTTTCAGNCTTGANAGATNANTNGTAGTTCTAGAAAAANANATAAAGTTTACTGNNAGAA  
NGTGGGAGGGAAGGACGGCNTGGC

SEQ ID NO: 37 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGANACANAGTCTTGCTCAGTTGCT  
CGGTTGCCAGGCTGGAATGCAGNGGCACAATCGCAGCTCACTGCAGCCTTAACCTNTGCGGCTC  
AAACGATCCTCCCATCTGTTTTTATTCTGTAAANATGGTGTCTCACCATGTTGCCTGGGCTGATCTC  
AAACTCCNGGGCTCAAGTAATCCTTCCTCTGCTTCCCTAAATGGTTGGATTACAGGTGGGAGT  
CACTCTGCCTGCCCTGNCAAGTCTTTTNCATNAAAACTTTTTATGTTTTTTTTTAAAT

SEQ ID NO: 38 TGAGAGGAAGTTCCATCGCCTAGGTTCTGGGAGAAGCAATACGTCACAATCC  
CCACTAAGGAGAGGGCTCAGGCAAAAGAGGAGAGTGACATTGCCTAGGGCATGGGCCCAGAGTT  
ATATCACAATGAACATGTGGACAGGGCCAACGCAGAAGTGTTAAATGACCTGTGTGCTGGGCCCCA  
GAGATATGTCACAGTTACTTTAATGGGCAAAGCCCACATCAGATAAAAGAGGCATCTCAAATATG  
TTGTAAGTTGATGAGCCCTGAGATATGTCACAATGTCCCCCCCCGAACAGATTCAAGGCAGGAGAT  
TCACATCACCTCGGTGCTGAGCCCAATGAAATGGCACAGTGTCTCCTGAGTGCAGGGCCAAGGCA  
AAAGAGCAACACCAACTTGGTATTGAGGCCAACCATATGCCACAATCCACTCTGGGAGCAGTTGT  
CAGGCAGGAGAAAAGAGTCACAACACCTGGGTTATGGCCCAATACATATGTTACAATCTTGCCCA  
TGGGCAAAGCCCAGGTTGAGACAGGAGAATCACATTNAA

SEQ ID NO: 39 ACTTCAGCCTGGTGACAGAGGGAGAGTCCATCTCAAAAAAGAAAAA  
ATGAGTAAGGCCTTGTGAGGAGCTCAATTCTTATCCTTTTCCCCAATGGTTGTGGATGTAGAGTT  
GTGCTCATTACCATCTTCTACTCAATTTTTTTTTTCTCATCCATAGTGAAATCTTAATCTTCAAAT  
ATCATGAACCTAGTTTTTCCACACTGGCCAGCCCATCTTTGCCTTTTCTCTCACGCTTTCAACTAA  
TGTAACAATTAGTCTGGCTATGGTCTAGTTTAAGGAAGTAATTTGAACACGAATCCTCCAAAGTG

GCTACATTTGTTCTTTCTACTGTCACATATAAGAGTGATGAAGTGCATTGATTTTAAAAGCTGGTG  
AAATTTGGGGGCATAGTTCCATGTCAGTAAAAAAATCCATGACTAAGTAGATTAAAATTATGA  
ATAAGAAAAATAGCAACACCGTATTTTAAAATCTACTTCCTATGCCTCATTGCTCATTAAATCAGGG  
AAATTATCAGTAGATATTATAAGTGAGGNATCTTATATTGGTTACTTA

SEQ ID NO: 40 ACGCGGGCAACTACGCTAAAGAATTTTGAGAACACCAGTGTGTCTACATTCA  
GTTTGCTGTCGTGGTTAGGATTCTTTTCAGCAATAATTCTGCCTTGACTAGCCAGTCTACCTTTTCCT  
GTGCTTGCACTCTTAATGCTTTGTTTTTTGTTTTTAAAGGAGGAAGGCTAAAATGTCAATTACTTT  
ATCCCTACCAAAGTTCATCTTATGCCTCTTCAAATCTACATAGCTTATGTCTCCAACTTATGTTCTC  
CCATTCAACTAATTAATAATAAGCATTTTGTTCCTTTAGCTAATTTAAGTTTTCAAATAA  
GGCCACAACAACAGGTCTCTGACAATCTCCAAATATCCTTGGGTTTATCACATCATCCTATATTTTC  
CAACAGGGACTTGGGCTCTACCAAGTATTTCAGTATAAATCTTTGTAAAGTAAAACATGGCCGGGT  
GCAGNGGCTCATGCCTTGTAATCCAGCATTTTNGGAGGCTGATATGGGCAGATCGCTTGAGCCCA  
CNGAGTCAAGACCAGCCTGGGCAAAATGGGTGAAACCCCAT

SEQ ID NO: 41 ACTAAGGTTACAGCTGTTCTGTTGGTCTTAGGCTCTGAGTAGACAGAGCCAA  
GATACTGCAGTCACTGGGATGGAAAGATGGAGTGCCTCCTTGGCAGTTTGTTCATGGGGTTAGA  
AGTTGTAGCTGCTCAGCTGCAGAATGGTGTGCTGCCATTAGTAGGTGTGGTGTAGTGGCAGTGAA  
GCCTAGGGTATGGGAAGATACAGTGGCTATAGACCCCCAAATGAGAAGGCACCCTAGCAGTGGCT  
TCAGTCTCAAGATGCCATTACACAGCAGCAGCTTGGATAATAGGGCAGGAGGAGACACAATGTGG  
GCTCCTTGTTTGGAGTAACATAGTTCATGTGAACCTCCAGGCAACCCCTCAGGCTGGCCCTAAGGACC  
TGTGAGGACTACAGTGATCTCCATGAGCCAAAGATTGTGGGTGTCCACATTTTAATTTTGATTGTC  
TGAAAGGCCTTCCTGCATACCTTTTCTNNTTGTAAAGGAGAGTCCGCCTTGGCTCTTGACCTNATNCC  
ACTAGGANAGACNAGATGGTNAAAGCAAATTGTTTCATTCCCTTTTTTAT

SEQ ID NO: 42 GAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATC  
TAGCATAATTTCCCCTTAGCGTGGNCNCGGCCGATGTACTGAANTATACTNGTCCNATGCTACAGG  
AATCTTTGGAATTTTATTACTATGNTTNTTCTAAGAAGAGGTATGNACCAA

SEQ ID NO: 43 ACGCGGGGGACTGAGAACAGGGACAGGCGACCCGACCCCAAGGGCCCGGTG  
CTCAGGACAGAGTAAAAGGCCAAGCTATGATAGCAACTGGTGGAGTGATAACTGGCCTGGCCGCC  
TTGAAAAGGCAAGACTCTGCCAGATCACAGCAGCATGTCAACCTCAGCCCGTCTCCTGCTACCCA  
AGAGAAGAAGCCCATCAGGCGCCGCGCCCGGGCAGATGTTGTGGTTGTTCTGTTGCAAAAATCCGGC  
TTTATCCCCATCTGGTTTTTTTCTTATTTTAGGAGTGCTCATCTCCATTATAGGAATTGCTATGGCC  
GTTCTTGGATATTGGCCCAAAAAGAACATTTTATTGATGCTGAAACAACACTGTNAACAAATGAA  
ACTCANGTCATTGGAATGAAGCGGTGTGGTGGTNTCTTTTTTANATATTTGCATTCT

SEQ ID NO: 44 ACTGGTGTGGAGTGAAGCAGGGCCACTTCTATGGAGAGACTGCAGCCGTCTA  
TGTGGCAGTTGAAGAGAGGAAGGCAGCGGGTGAGTCTCCAAGGACAGGGCCTGCACCCCTCAGA  
CCCAGAGGCAGGACTTCCTGAAGGCCCTGCCTGAGAGCTTCTCAATCAGTGCTGGCCCTATGTT  
TGGCTGTAAGAGGCTGAAAGTGGAGAGTGGGAAGGGAGGGGACATTTAGGTCCTATATAGCCTCG  
TTGAGCCCTTCAAAGGGACATCTCATATAAACATACCAACTAATTAATAAGTGGGTTTGCTATT  
TAACCTCGGCATTNAGACAACTCCCACTGAATGAGTGGCTCACACCTGTATCCCATCACTTTGGGA  
GGCCGANGTAGGCANATCACTTAGGCCNNGGANTTCAACACATCCTGNTGACATGGAGAAAACT

SEQ ID NO: 45 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGNTAGTTTCTATGACTATGTCTAACAG  
TTAGTAATCTTTTCTTCTGCAATGCCTAATTGCCTTAATCCCCTCTGGTGGTGGTTTTGTTGTTGT  
TGTGTTGTAGATGAAAAAN

SEQ ID NO: 46 ACGCGGGATATGCTTGCAAATTCAATTTAGTTAAATTAACACAGTCTTTAAAA  
TCTCTAAATATTGATCTCCAGGTTTGAAGAACTTTCTCTAAATAATATTTATGCCAGAATTTGGGCA  
GCCAATTCATAGTCATGATATCTTCAAATATCCTAACCTTATCANCAAAATCTTTGTTCTTCCAAA  
TTTAGCTACTTATTTAAAACTACATAATGTCTTTTTTTTCTTTTCTTTGCAAGNTNACATAATGT  
NTTTTTTNTCTGNGCTACNAAGATGTTTNCCTATNTATGAAAGGNNCTANATTATTGCCAGTTGC  
GGTGGNTTA

SEQ ID NO: 47 ATGNGGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGA  
TATCTGNTGGAATATCCGCCTTTCGAGCGGNCGCCCGGGCAGGTACAAAAGGAACCTTGACATTT  
CNCAAAAAGCATTTCACTTGAACCTCTGCTTAAAAAACTGNTTCTNANCANCGTNATGAANACA  
AACCANTAAATGTTAATNAAANCTACAAGATTTATGGCTCTGAGAGAAATATACTGANTGATGCA  
TCNTAANTATCCACAAATACCNATTAATGNAATGTTTAATACTATATNAT

SEQ ID NO: 48 ACAAGTAGAGAATGTCTTTACTTTTTTCCAACCACTGCCATCTCTTACTCTGAT  
GGATTTTTCTTCCCAGCCCTCAGATGGAAATGCAGAAATCACCTGTCTTCTGCTTCGCTCACGCTG  
GGAGCTGTAGACCGGAGCTGTTCTGTTCGGCCATCTTGGCTCCTCCCTGCCAGATAGCTCTTATTT  
TGAGCTCTGAAGGCTGAGAACTTCAAGATCAAGGCACCAGTAGACTCAGTGTCTTATGAGTGCTA  
CTATTTGCTTCAAAGACGATGCCTTCTTGCTAAATCCTCTGCAGAGAGTAACAGTGCAATTTGACA  
TAAACAGCTCCTTGCCTATCTGGCTGCCAGCCAAGCTGGGTCTACACAAGTTGCTATAAGAGCT  
TTCTACAACAAGATCTGACTACTGAAAGATTGAGATCACTTGACTTTTAGCTTGAAGAGAACATTT  
TATTAGTGGATAGCGTTGTTGGAATTCAAAAACAACATATTGCATAGAAATGNGTGAAGATTGCA  
TTACCTTATGGATGATAACAAAATGCAAATCTCAAGGNATTTTAAACAGTGTNAATCCAAATTGTA  
CTGAACTCAGGNGTTTCTAATGTTCTATGGATTACTTGCGCAGTGAAGGAGG

SEQ ID NO: 49 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGCCAATATTTAAAAATATAATTG  
ATATTTGTATGTTGACCTTGCATCCTGCATNTTGTCTAACTCACTTTAGTTCCAATAGTTTTCTAT  
ANATTTTAGGGGGATTTTCTATGCAAACAATCATGTTGTCTGTGAATANACATGGCTTCATTTATTT  
TCAATCTGTATGTCTTTTATTTCTTTTGTCTTACTGCAANGGCTAGGACCTGTAGTGCAATTTTGAA  
TATGAGNGGGGAAAGNGGACATCCTCAGTTTGTCTTAATCTTACAAGGGGAGCATTTAGTCATTC  
ACCTTTGAATATAAAGTTAGCTGCAGCTTTTTTNGGATCCCATTTATGAGGTTGAGGAAATTTCT  
TCTAAATTNGTAAGAGTTTTGATCATCAACAAATGTAAATCTATCAAATGGTTTTCTGATCTATTG  
ANAGANCAAATGCTTTTCTTCTTATCTGN

SEQ ID NO: 50 ACATTTACATTCTGTAAGAGATTGAGCCTGAACTCTCTTAGTCATAAAAAACAT  
CAAATGGCCACATGTCCACTACCAAGCTTCTTCTATGTTAAAAAATAATAAAGCAGTTTTAA  
CCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 51 ACGCGGGGTATGGGGTTTCTTTTTGAGGTGATGGAAATGTTCTGGAATTAGAT  
AATGGTAATGGTTGTGCTATTTTCATGATTATACTAAAACTGACTTTTAATAATCCACTAAAGTGA  
ATTGTATACTTTAAATGTTGAATTTATGCTATGTGAATTATATCTCAACTTTTTTAAAAAAGAGG  
CAACATCAGAATCCTAGAATTGGAAGAGNCCCTGAGAATTGTGTGGCCCAAACCTTCAAGTCTTG  
GCAGTTGAGAAATTTAAGGGTATTGCCAAGAAATNTGTCTTTAAAAATAAATTTGAAGCAGGTTT  
CTCTCTTGCTTAGAACCCAAAGAGACTTGAGAAGGACAGCTGGCNTTAAAAACCTTATGATCCCC  
AANACAGGGTGGCCANACCATGCANACAGTNGCCTGTCAACATNTGAGACCCTTCTTCATAAAATTG  
AAAGAANAATACCTGACATGGNNCAACTGCCTATATNCCGCTCT

SEQ ID NO: 52 ACGCGGACACAGGCAGTCACTAAAGGGATGGCAAAGACAGAAAGAAATCTT  
ACTGTTTTATGTAACATAAATGGTTACAATCCATTACATCCATATTTTCAAGATAAAATTGCTAAITTT  
TTTTTCACATTTTATGACTAGAGGTAGGTTTTACATTTTGGAGTCAGGTGACAGATTAGGCCCATC  
TCCTCCTAGGAAACCAGGCGATTGAGTGTTATTTCCCATTTTCAGAGATGGCTTTCAGGTCTTAGG  
AAACATAGCCAAAGTTATAAGACTAGAGACTTAAATTTATATACATTGAAAGGGAGGGAGAAATAA  
AATCTGGAAAAAGAAAGGGGAAAGGGACCTTCCCTTTTATTTTCAGCAGGGAGAACTATGCCT  
CTTATTTTAAATTTGATTTGCCTTACGTCAACAGCAGTATTTTGATCTATACCATGGCTGCCTAAGT  
GATGTAGCGTGATTGCAGTAAATATGCCTTGATCTCACATACACTTTGTCTAGG

SEQ ID NO: 53 ACTNTGTGATCTTGCTGAAGACTACAGGCAGCCAANTGGTTCCAGATACTTC  
AGCTTTGTGTATCTTCTNAACTTNATATTAATATAAGNTTCTNAAAGAAANCCCAAGTTTTTAATCGT  
TNTTGATTTAAGGANAAAAATAATCANAATGAATTTNTTGCATAAACNTTAAAGTCA

SEQ ID NO: 54 ACGCGGGGAGGCTAGCCAGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTT  
TGGGAGGCCAGGGTGGGTGGATCACTGAGGTGAGGAGTTTGAGACCAGCCTGGCTAACAAGGTG  
TAACCCCATGTCTACAAAAAAAAAAAAAATTAGCCAGGTGTGGTGGTGGNCACCTGTAATCCNAGN  
TACTGANAAAGCTGAGGCAAGAGAATTGTNTNGAACTGGGGAGGNGGCTNNACCAGGGNAGGCA  
NANGTTGAA

SEQ ID NO: 55 ACCACTTTGGAAATGCACTGACTCTTTAAAAGCCACATAAATGTTTCAGCCATT  
AAATTCAAAATNTTATGNCTTAATAAAAAATGATTCCTCCCATCTCAAATCATTATTTCTTTCAA  
CTNTCATNTNCNGNAACCTGNTNCNCAANCCNTACNTTGNNTTATGGGAAGNTTTTGAACCNNTN  
TTTTATCTTCATTCCTTGGTNGGAAAAANAAAAATCTATTATTGAATCTTTTACACACTNNCNTTAA  
ACNNTTNGATTTTCTTATCCCTTTTACTTTTCAANAATTTTAAACNNGGTTATCCCTACATTTGG  
GATNATACCACCCCTTGGGAAGGGGGGNGCNTTGGCCANCCNAAANCAACCNATGGGGNANG  
GGCAGNCTA

SEQ ID NO: 56 ACCAGCGGATCGCANC GTTGCCATGAAATCTGTGCATTTC AAGGCCCGTTTGG  
GGATGTGGGTAACTCAGTGGTTCTCAGGTGGGGGAGGTATCATGATTTACCTCCAGGGAATACT  
TGGCCATGTTTGGAGACCTTTTTTGGTTGTACAGCTAGGGGAGGGGTGCTATTGGCATCTGGTGG  
GTCAAGGCTAGGGATGCCGNTGAATATTCTACANTCCACAGGACCACCCGCCNCAAAGAAGGNT  
CCCANCTGATATGTNAGGCTTGCCACNGNGGGGAAACCCCTNATTAAAGTATTAAAAATTTGAAN  
CTACATNTTNTTANTNCCNCGGGGCTTTTGTTTTTA

SEQ ID NO: 57 ACAACCACTATGGGGAATAGTTTGGAGGTTTCTCAAAAACTAAAAATAGAG  
CTACCATAAGATCCAGCAATCCCACTGCTGGGTATACACCCAAAAGAAAGGAAATCAGTTTATCT  
AAGAGATATCTGCACTCCCATGTTTATTGCCAACACTATTACAAATAGCCAAGATATGGAAGCAAC  
CTAAGTGTCCATCAACACATGAACAGATAAAGGATATGTGGTATATATACGCAATGAAGTATTAT  
TCAACCATAAAAAAGAATGAGATCCTGACATGAGGTCATTATGTTAAGCAAAATAAGCCAGGCAC  
AGAAAGACAAACATCACATGCTCTCACTTATCTGTGGGAGCTAAATTAACAATTGAAATCATG  
GAAATAGAGAGTAGAAGGATAGTTCTAGAGGTTAAAAGTNTCTCGGCCGGACCCNTTAAGGCNATT  
CCACNCCTGCGCCGTACTATGGTCCACTCGGACAANTGGGGNATATGGATACTGTTCTGGGAAT  
GTNTCCGTC

SEQ ID NO: 58 ACGGCCAGGGCTATTTNTTGAATGAGTAGGCTGATGGTTTCGATAATAACTA  
GTATGGGGATAAGGGGTGTAGGTGTCCTTGTGGNAAANAAGTGGNCTAGGGCATTTTAAATCTTN  
NANCGGAAAGCNTATANTCACTTGCNCCCGCTCATAAGGGNTTGNCTTGGCNNGGTTTATATAT  
AGTNGGGGGGTTGCGTGTAATTNAATGA

SEQ ID NO: 59 ACCTAGAAGAGAGGCGNTTCAAAGAAGTAGTGAAGAAGCATTCTCAGNNCA  
TANGCTATCCCATCNCCCTTNTTNGGAGAAGGAACGAGANAAGGAAATTANNGATGATGAGGCT  
GAGGAANAGAAAGGTGAGAAAAATGAGGTAAATCNTTATTGATTGATGAAAAANCCAAAAATC

SEQ ID NO: 60 CGCGGCGAGCTATCNTTTGAATANTGAGACAGAAATNAATCAATATAGAGGC  
TGTGCACGGTGGATCACGCCTGTAATCCAGCACTTTTGGGAGGCCANAGGCAGGTGGATCGAGA  
CCATNCTGGCTAACATGGTNAAACCCGGTCTTTACTAAAAAATACAAAAATTTNCTGCNTGNGGG  
TACCGGNCACCNAT

SEQ ID NO: 61 ACGCGGGATATCAATAATGGGTCTGATATAGACTGAGGATTCATATTAAGTC  
CACATGCCTCCAAAAAGGCAACCTAGAGTCATGACTAATACATGGAAATTTGGTGCCTCCACCCGC  
AGCTGACCCCTTTGGTCTCTTAAGAAAAGAACTAGAACTTTTTAAGGTCTGAGATCAAGATCTTAC  
TTTTTTTGTTTAGTAAGTATTTAGCAAATATTTTTGAAATAATTTTCCATGAGAAGCATGAACATGA  
GCTACATGTTTGAAGTAAGGATGTAATTGTAGCTTCCACTTGCCTNTCAACATGGAAATGCTAGAAG  
TTTTACTTACAGGGTTCAAAAACATGTATACAGTCATCCCTCTGTATCTGNGAAGGATTGATTCCA  
GGACCCCTTACGGATACCAAAATCTGNANATGCTNAAAGTCTTTGACATAAAATTTGGCATNTNNTT  
TNNCATAATNNACTTATGCNNNTCTCCTATATTACNNTANNATTTTNTCTACATTACTTATTTTTACC  
AAANNCAATGTAAATAGTTCTTTAACTGGCATTGGNTTAAGGGGAACNAC

SEQ ID NO: 62 ACGGGGAGACTGTGGAGCANTTATTCAAACCTCGGAGGGAGTCGGCATGG  
GAGGATCCATATAATTTACGCTAAATTGTGCNCGTCTGTTTGTGAAATGTGAAGGNGCACATTG  
TTTTCTGGAAGGCAAATTTCAATTTNTTATACCACTTGCCAGAAAGATCTGTGATCCCAANGAAC  
TGCTGTTGTTANAAACAANGACAATCATTTTGANCAANAAATGATGGTTCCAACNAGGGAGGGA  
GTAACCATGGATATTGCTGAAATGCAGTTGGTGCCAGGGATTTATTANGACATGATTAGTTCTGNA  
ATCATCCCTAANGTAGCGATGAAGTCTCNCTATGTTGCCAGNCTGATCTCAAACCTCCGGCTCTA  
AGTGATCCTCCACCTCATCACTCCCAAANGTGCTGGGAATTAAGGCCTGANCCCATTTGTNGCC  
CAAACCTNAC

SEQ ID NO: 63 GGGGACTGAGAACAGGGACAGGCGACCCGACCCCAAGGGCCCGGTGCTCAG  
GACAGAGTAAAAGGCCAAGCTATGATAGCAACTGGTGGAGTGATACTGGCCTGGCCGCCTTGAA  
AAGGCAAGACTCTGCCAGATCACAGCAGCATGTCAACCTCAGCCCGTCTCCTGCTACCCAAGAGA  
AGAAGCCCATCAGGCGCCGGCCCCGGGCAGATGTTGTGGTTGTTTCGTGGCAAAATCCGGCTTTATT  
CCCCATCTGGTTTTTTTCTTATTTTAGGAGTGCTCATCTCCATTATAGGAATTGCTATGGCCGTTCTT  
GGATATTGGCCCCAAAAAGAACATTTTATTGATGCTGAAACAACACTGTCAACAAATGAAACTCA  
GGTCATTGCGGAATGAANGCGGTGTGGTGGGTTTCGCTTCTTTGAGCACATTTCATTCTGATAAGAT  
GAAAATGCTTGCCATTCAACATGGGGATTGGCATTTTTCATTTTCATTTGTGCTAATGCCATCTTCA  
TGAAAACCGTGACAAAGAGACCAAAATCATACCATGAGGGATATCTTTCCAGTCATTGACATTC  
ACACGCTAAGAATAAGGAGCAAAGCAAATG

SEQ ID NO: 64 ACGCGGGGCGGTCGAAAAAGAGATAAAGTTGAAGGAAATAAAATTGGCACG  
GAGTCTGGGAAAGTAGTTTCCCTAAAGGAGTCTTAGAAATAGGGTTGGTCTGGAACCTAAGGGGC  
GGAGCCGACGCGTAGAGCCGCTTTGCGCGTGCATCACCTAGGCGGTTAGATTTGAATACTTCAC  
TGAGGCGAGCCGGGCGTTGNGAGCGGACTGCTAGAGGCGGCTGTCTGTTCCGCTCTAAGGAAAC  
TCAGAGCGTGTGGACCCCAAACAANTCTGCGCAAAATTTGTCGAGGAGGTTGCCNCGGCAGAAA  
AGATTTCTTCAAAAATGGATGGGGTNGCCTTCAGANGCTTATTAAGAANNNTCTNTTTATNNAGAC  
CCCGTTGAAGACGGCTTTNTTCANTATTNANCCCNAAAGGAGTTCTCTCATGACCTCTAAGGTGGAA  
TGAACANTTCANGAACCNATGCTTGAGAAAAAGAAAANTCTNTCGAGCAGCTTTCNGATCTANAA  
GNTTCCAACGAGTNTTATAAATAC

SEQ ID NO: 65 ACGCGGGCACCACGATGAAAGGGCACTGGCAATGGGAATGGCATCTATAGT  
GTTGGCATTCTATATTTACTTTTATTATCATTTAGAAGGATCCTTCTAATCAATTTAATAGACAAAT  
ATTTGCTGAACACATAGTAGATAACAAGGTATTGTGCTGTGGGGTTGTGAGGGTAACAGCGTGTCT  
TCTCCTTAATAACAAGCTTAATANCATTAAATAGNGTGAATTACTATTTAGAAGGATTAGAAATATT  
AACNCAATGGGAGGCTGAGACACAANAATCACTTGAACCCAGCANACGGAGGTGGCAGTGAGCT  
GAGATCTGCACTCCAGCCTGGGCAACAAGAGCAAACTCTGTCTCCAAAAAAAAAAAAAAAAAAAA  
GTCCTGCCGGGCGGCCGCTCGAAAGGGCGAAATTCATCCACACTGGCGGGCCGNTACTANTGG  
ANTCCAAGCTTCNGTACCAAACNTTTGGNTANATANATGGNCCATAGCTNGTTTCCCTGTGGNNA  
AATNGTATCCNCCNNAATTTCTCACAACANTACTAGCCCGGAA

SEQ ID NO: 66 GGGACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGANACANAGTCTCACTNTGTT  
GCCAGGCTANAGTGCAGNGGCGTGGTTTNTGCTCACTGNAACCTCCATNTCCAGGTTCAAGCA  
ATTCTCCTGCCTCAGCCTCCCAAGTAGCTGANATTACAGGCACCTGCCACCACGCTGGCCAATTT  
TTGTNTTTTAGTANANACAGCGTTTCACTATNTTGGCCACNCTGGTCATGAACGTTNTGCCCTCAG  
GNGATCTGCCCNCTNANCCTCCCAATGNGCTGGGATTATAGGCGTGAGCCACCGCACCTGGCCC  
ANTTAACCTTCAAAAATGATAATGATCATGGCTCAATTTGGGGTGATACTTGNCTCGGGGCTCATT  
GACATTTNTNGGTAAACNCANGCTTTGNTGTGACAGCCTGGAAATGATACTGCAATTTGGCCCCACA  
CAGGAAAAATTTTNTATTTTNCACAAAANACTTTTGTGTTGTTGGTTTTTTGANAAAAANCTN  
GCNTTTATTGGTCTTGACTGGGAGTGCAAGGGGNAATGATCTNAANNTT

SEQ ID NO: 67 ACTGAAGTCAAAAACAGCACATGGGCCTTGACGATCTGGGGTGACGCAAGC  
AGGGTGAAGGGGGAGACTGCCTTAGAATGGAGGGTGGCAGCTCCAAGGAATGGGAAAGCTTCTC  
AATCCCTTTGGTCTAAGTGGGAAAAACAAAGAGCTGGCAGCAATCTGGGGTGAGTTAGGGCCCAG  
GAGCTAAAAAGCAGGCGGTTTCATGATAACGAGCTTCTTTTTTGTGTAGGATGACTGTAGCAAGT  
TTATCTGCAAGACCCACAGGGACCTCAGCAATGACACATTGTGGGATAGATTGGGAAGCCAGATT  
AGGTAATCACAGAGCTGGAGCACCTCTCTAAGAACCAGCTGGGGCTCTGCAGGGGTGTGACAAGT  
GCTCATCCGGATGAGGGACTCACAGAGACAACTGGCACATTAAACAGATTGCACTGTCTTCTCCT  
GACAGCACGCCCACAAAGGACCATGCTCAGCTGTCTCTTCAAAGTGTGGGAGCAGCTTCCCCC  
AACCCTGGCANGAGGCCACAAGATCCAAGTNGGGACCCAGCCTCNACAAGNG

SEQ ID NO: 68 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGACGGAGTCTCACCGTGTGCCA  
GGCTGGAGTGACGTGACACGATCTCAGCTNACTGCAACCTCCGCCTCCCGGTTCAAGCGATTCTCC  
TGCCTCANCCTCCCGAGTAGCTGGGACCACAGGCGCACATNACCATGCCCGGCTAATTTTTGTNTT  
TTCGTANANACGGGGATTCACCGTGTANCCAGGATGGTCTTGAACCTCTGACATNACGCAGTCCA  
CCCGCCTTGGGCTCCCAAAGTGCTGGGATTACAGGCTGGAGCCNCTGCACTTGGCCATTTTTTTTTT  
TTTTTCTGAAACAAAGTNTNACCCTNTTGGCCNNGCTNGAGNGCAGTGACAATNATNTCGGGNT  
TANTGCAACCTTTGGCNCCTGGGTTNAAANCGANTTTTCTGCCTAACCCTTGAANTANCTGGGG  
ATTACANGCTCAATGCCNCCACCCCAAGTAACTTTTGTNTTNTNCCNCCANNTTTNGNTCAAGG  
CTGGTNTTTNAACTCCTNGCCCTGGAANTTNCCCCACCTTTGGC

SEQ ID NO: 69 ACTTAAAGTAATGGTGATCCTTATTCCAGGGCTTCGCCGCCAGGATTTCTTGC  
TGCTGTTTCTTGTGCTATTTGTANATTTTCATCGATACTCTGAGCTTCTGCTGATGTTGTTGGCCAGNT  
CCCANGGTTANTCTGGGCCATACCGNTNCACNGACGCCGTCTTCGATCGAGGAAACNGA

SEQ ID NO: 70 ACCACCTTCTTAACACAAATGATTTAATTTAACCATTAAAGTCAAGTCTGCAAT  
ATCAATATCATCTTGCTCATTTTACAAGTCACAAAAATGAAGTTCAGCAACCTTAGTTTTGGAGCT  
GGTATTTAAACCTGAGGCTCGTCAACGTTAAGTCTGTGTTCTGCCCCAACAGGAGTATAGTCTTA  
CAAAGACATCTTTCTCTTCCAGTCTCTGAATATTCTGAATTATCACAAGTTAAGTGACTTTGATGT  
GGTCTTCTCTCCCTTCTCTTTCATCCATCTATCCATCTTTTCATTCAATTTAGTGATGTCCCTTTAT  
GCACTGCAGTGAAAAATAAAATGACTTTTACTCCCACCATACTGCAGGGAGCTATGACTCATATTT  
TTAGAAACTGATCTTCTGGGACTTTTCCACCTGTATATATTTCTCCAACCAGCTTGGATTTTGAA  
TAAGGAACCAAGGTTTTTTGGCTTACTATTGAACATTTTAAATCATAAACATTGGATTCTCTACA



TAATGGATTCTAATTATTTAAATAAAAAGCCTGGGCCGATTATTTGGATGAAAACATATTTCTATTTT  
ATTTTCCCAAGGAATGGNAAAGAATATACTGGAATAAAAAACAAAATCCCCTAAGGGAATAAT  
AGCAGGGTAAATTTTATTGGAGGGGACATGCCGAANAAGGTGTTATTTGGGCTTCCA

SEQ ID NO: 71 GGCATGCTCNAGCGGCCGCCANAGTGATGGATATCTGCANAATTCGCCCTTT  
CNAGCGGCCGCCCGGNCAGGTACGCGGGGCCAGGTCTCTNACTTCTGGCTTGTTCGCTGGTGGCG  
GTCNNANCCGAGCCGACTGGTCANAGATGATCACGGACGTTTCAGTTCGCCATCTTNNCCANCAN  
GCTGGGNAGTGTGNTNTTCTTGCTTGNCGTTCTCTATC

SEQ ID NO: 72 ACTTTTTTTTTTTTTTTTTTTTTTTGAAAANATGAGGTTTTGCCATATTGCCCAGG  
CTGGTCTCAAACCTCCCGAGCTCAAGTGATCAGCTTGCCTTGGCCTCCCAAAGTGCTAGCATTATAG  
GAATGACCATGGCACCCAGCGGTAAATGTTTCTTTTCANACTTTTAAAGGTGTCAAACCTGGCTGG  
GCNGTGGTGGCTCATGCCTAGTAATCCCANCAANTTTGGGAGGCANANGTGGGCANATCTGNTCCC  
CTTNTCTCTTCCAAGTTTTT

SEQ ID NO: 73 ATCTTTATATTATTTNCTTAAATTGATTGGGCCCTCTAGATGCATGCTCGAGC  
GGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTAGCGTGGTTCGCGGCACGANGTACGCGG  
GGGTAGGGAGGGGGACCAAGTGGCAGAGGGACCTTAGGTGATCCTTANAAATAAAGGCTAGTTTCT  
GTTTCGACCTTGGAGTANGCGGAANAGGTGTANACAGGTCTGGAGAANCAGAGGTAAACCTGAG  
TAAAGCAAGAAAGTTGGAGAATATGAGATACATCTCATCTCTTTAAATACTTAAATGACTTCCCCT  
CCTCCCGGAGTTNTATCAAAATTCGGNGATNNANNTGACNGACGTANGTGAANACNCTGTTGGGA  
ACTTACANACTAAACTTG

SEQ ID NO: 74 ACCAGCCCAGAGAGGCTCTCTGCTACCTGACTTTCACTACTCTATGGTAATGT  
GCAATTTCTCCCGCAACTGAACTACAACAGAAGTTTAAATGTCTAGCCTACAGATAGATGTTTCCA  
CATGCCTAAACTCCACAGGAATGAGTTGTCTTTTACTATGTGAGAAGTCAAATGTAATGTTGGCAA  
ATACCAGTGTGAGAGCTACAAGTTTCAACAAAAGCAGCACACTGTATATTGGCAGTCAAAATCAC  
TTCGCTCTATTCTGCATTAGGTACGCCAAACGCTTCTGTCAAGACTCAGTAACTCTCACCATACAT  
TTTGTCTATTTCATTGGTCTATGGGCAGGAATGTCTTATAAAGCACACATTACTAAGTGCTTTATAT  
ACATCACCTTATTCCTTAAAAAAATTCTATAAGAAGTATTATTNTCATTTAGAGATTAAAAATTGA  
AGCTCCCGAGGCTTACCAATTTCAACTAATAACTAATTAACAGGGCTGGGATTGTAATTCAGGCT  
ACCTNCTTTAAATNTGGTTTTTTCATTGGTTAGG

SEQ ID NO: 75 ACTNTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGGGTGTTGTTTCACTTTCTTGGN  
ATTGNCTTTTGATGCATAATCATTTTCAATTTTGAGGAAGTCCAATATATATACTTTTTCTTTGGTG  
CTTGNCTTTTGNGTTCATATCTAAAAAACCATTTGTCTATTCCAAGGNCATGAANATTTATACTCA  
TGTTTTCTTCGAGTTTATACCTTTAGCTTACATTTAAGTTTTTGATGCNTTTANAGTTAATTTTTC  
CACATGATANAAAGTAGGGATTCAAGTTTNTTCTTTNGCAAATGAATATCCACGTGCCCCAACATC  
ATTTGGNGAAAACCATNCNTTTTTCATTTGAGTGAGCGGGGCNCNCTGTNGAAACAGAAAANC  
ATATATCTATGGGGTAATTTCTTGGACTTTTGCTTTATTCANTGATCTACNGGGCTATTCTTTNCA  
AAAATNNCNCNGNTTNTNCTGGANCTTACTAAGTTTAAANCNAGAAGGGNAAACCTCCACTTTG  
TTTTCAATTAAGAAAAATNTGG

SEQ ID NO: 76 TTTTCAATTTTGTANGCATTGGGCCCTCTAAAGCATGCTCGAGCGGCCGCCAGC  
GGGATGGATATCTGCAGAAATTCGCCCTTAGCGTGGTCNCGGCCGAGGTACTCAATTTCTGTGGAA  
CTNTGNTATCCTGNGGAACACACAACATTACAGTGGAANCTTCTGTGAAAGATGCCAATGGTNTA  
GTATGGAAGACTGGTCCCAGAACTACCATATTTGTAAATCCCTGGAAGACCTTATTCNCAGCAA  
ATTCNNTTACAACAGCTNGTAANCCCANCATGNCATTACANNGAGCANCAGCTACGGCCNNTGCC  
TACACACGGTTTTTAACCANNGGAATGAATGCNCTTGCTGNACTTCTTAAGAGATCAAATCNAATTT  
GCCNTAAAACAANTTTGNNATTTTNNATGGAATAAGGTTTTANAGTTTACA

SEQ ID NO: 77 ACTTAAATGTAGTAGATTCTATGCCTATGCATATTTCCCAAAATGTAAAGTG  
AGAATTGGAAATGCAACTCCCAATCTTCAAATCTGAAAAATTATTCAATCTCAAAACTCATATTC  
ATACTATATAATAACTTATACCTCTTCCCTCTGCGACTTTCACATATCAGCTGGAGTTGGGGG  
AGAGGAGACTAAAGCACCCCTCCAGGCACAAATTTACCACAAGTTCCCAATAAACCTTTACCAAC  
TCACTATGACAAAGCATAGATGAGAGAAATAACAGCATGCCAGGTCAACCGAGTCCATAGGTGA  
GTGTAGCTGCTCANTAAGTGTGTTGGTTGATTAAATTAAGGGTGACTNAATCCATGCCCAAACTGAG  
TCCTTACNAAATGCCCCATAAAATTTAAATTTTAGAAGAGTTAGTAAAGAGCTTCTTGAATACTAA  
CTGCATGGAGATACTACACAAAACAGTCATNTTAATTCCTACANCCTTCANAACAAAAGAGTCC  
AACTGAAAGTAAGATCCCTGTNATTATGTTGATCCTGGCT



SEQ ID NO: 78 ACCTTTTTTTTCTTATTTAAAGCACAAAGAGGCCCATAAATCTTGAGTTACTTTA  
AATTCTTTTTTTTGATACAAAGTTTTTCAGAGCAAGAGAATAAAAATCATGTGTTATTAAACCCCTAC  
AAAAAAAAAAAAAAAAAAAAAAAAAGTNCNCGGGGTCTTTGAAATAAAACTCCTTGTTTTGGCCGTG  
TTTCTGAATGTATTGGTGATTCACCCCAAGCTAATTTTTTAAAGTCATTTTTGAAGTTTGGGAAGT  
CTCATGAGAGATGTTGAANTATGTNTTTAATCAAGAGTCATGATTCAAACCTAGTTTTACATATTA  
AGCAGNTNGGGNTCTAATTTNATGGGNAAATGGGGGGTTTGGTNAAANNTCTGAAAAATTTNTC  
CCTTAANGATATNATAANTNATTTGCCNNTAATATGTCTTAAAGCTNTTTNAAACAAGGGTGTTNA  
TTGAGCCNAAAACANTTNATTNATTA AAAACAGTCCTGNCCGGGCGGNCGTTCNAAAGGGCG

SEQ ID NO: 79 ACTTTTTTTTTTTTTTTTTTTTTTTTNGANACATGGTCCCGTTCTGTCAACCCAGG  
CTGGAGCATAGTGGCACAACACGGTTCACTGTAACCTTAACCCCTCCAGGCTCAGGNGATCCTCC  
CACCTCANCTCCCAAGTAGCTGGGACTACAGGNGCNCACCACCATGCCTGGCTAACCTTTTCAA  
ACCCCTGGCCTCAAGNGACCCACTCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGACCC  
ATGGCGCCCAAGTCAAGAACTTTTTTTAAACAAGCCATTTAAGAGTGCCTGCTGCTTAAACAAAAA  
AAAAAGGGCNTACAATTTCAATTATGANCAAAATTTTGAGNCCTAAGTAAAAANAAGTAATTTTTNG  
TTTTTTTAAACCAGCTTACATTGTTNTTTAGAACAATGACTTATTTCAAACCTTGTTAAAAACA  
AAAACCTTGAGANANTAGAAAAAATNTTTTTTAAAGGATTAGCTTTAAAGNTAATNATAAAGAA  
AAAAGTTCNCTTTGTNTNAACTTTGAGGAATGGATTATAGGGGAGAA

SEQ ID NO: 80 ACGCGGGGATTCTGAAGCTGGCAGCATTCTGGGCCGAGATGTCTCGCTCCGT  
GGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTGGCCTGGAGGCTATCCAGCGTACTCGGGTTGAT  
TCCATTCCAATCCATTCCAATCCATGCCATCCACTCGTGTGATTCCATTCTTTCCATTCCATTCAA  
GTTGAATTCTTTCCATTGCCATTCCATTCCATTCCCTCCCGTGCAGTCGGGTGATTCCATTCCATTCC  
CATTCCATTCCATTCCATTCCATTCCATTCCATTCCCTTTCCGTCCGTTTCATTCCATTCCATTCCATT  
CTATTCCGGTTAAATCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCAATTACATTNC  
NTTCCATTCTATTCCGAACCTCGGCGCGACCACGCTAAGG

SEQ ID NO: 81 ACAGTTCCCATCACGTATGTCAGTTTTGTGTATGCAGCAGAAATGATACCTA  
AACTAAAAACAAGGACACATANAACAGGAGGTGCTGACCAAAGTCTTNAACANGGAGAGGGAAG  
TAAAAAAGGGAAAGGAAAGAAAAAGTAACCTATTATCAGCATCAAAGTATGTGGTACCTGC  
CCGGGCGGNCGNTCCAAAGGGCAN

SEQ ID NO: 82 ACGCGGGGACATAAAATNTNCTTTAACGCATTTAAATAAACAGAAATCATAC  
NAAGTATGTTTTAGACAGTAGCGTTAAATTAGAAATAAGTAACAAAAATATAGCTGGAAAACTC  
CAAATATTTGGAAATAAATATATCTAAATAACTCATGGGCCAAAGAGGGAGCTAGATGTAAAAAC  
AGAAAACTTTTTGAAGTGAATGAAAATAAGTGAACTTANATTATACCTTGGGAAACTTGAGAAA  
GAAGAGAAAATTTAAGCCTAAAGCAANGCAGAAGAAGGGGAATAATAAAAAATTAGAGCAGAAAT  
AAATGGAAATTGNNAACAGAAAAAATAGAGAAAATAANATGGAAATTAAANGTTAGGTATTTAT  
AAAGATCANCAGACTGGTAAACCTGTAGCCAGGCNCCAAAAAAAAAAAAAAAAAAAAAAG

SEQ ID NO: 83 ACAAGGAAAACACAAAATATGTATGAAAGAAATTGGAGATGACACAAACA  
AATGGAAAGACATCTTATGCTCACAGATCAGAATAATTAACATTGTTAAAAATTGTCATAATACCCA  
AAGAAATTTGCCGATTAAATGCTATTCTATCAAAATATCAACAGCATCTTTCCAGAAATTAGGAA  
AAATGATACTAAAATTCCTGTGGAACCAAAAAGGAGCCCAAATAGCCAAAGCAATTTGANCAAA  
AAGAACAAAACNAGAGGCATTACTTTCTTATTTCAAATTTTTTCAGATTCTATAGTACCAAAACAG  
CGTGGTTTTGGATTTAAAAACAGACNCTTCCCATGGAACAGAATAGATACCCAGAAATNATCCTTT  
ATTTAGCCAAACTGATTTTACAAAGGCATGAAGACATACATTAGGGAAAAGACCCCTCTTCAA  
TAAACGGTGTTAGGAAACTAGATTTCTTATCCAGAAGAATGAACTGAACTCCTATCTTCACTATN  
TACAAAAGCCAACTCAAGATGAGATTAAANGCTTAT

SEQ ID NO: 84 ACCGCCGGGCANGTACNCGGGACCATACCATATCCCACCAGAGAGTGACTCC  
TGATTGCCTCCTCAAGTCGCANACACTATGCTGCCTCCCATGGNCCTGCCCATGAGTNTNTTGAAT  
GCTGCTTC

SEQ ID NO: 85 ACGCGGGAGGCGTGAGCCACCGAGCCCGGCCACAATGTGTTTATATACACAA  
AGGAATATTATTCTGCCTTAAAAAATAAAGAAATCCTGACATATGTGACAATATGGATGAACCC  
GGAGGACATTATGTTAAGTGAAATAAAGAAGGCACAGAAATACAAATACTGCATTATTTCAATTA  
CAGCGTAGAATCTAAAAATAGTCAAAATCATAGANAAGTANCATGGTGATTTCCTCGCGTGTTA  
ATTACCAATCTTCCCATATAGGTAATTCCATATGGCACAACTGCCCTTACTTGNAGAGGCCACCNCA  
CCCTGTGNCATGAAGTGNAGGTGTNCTTCCATGGTTGCTTGCCCANTCAGGGCCTGTATCTCTTTC  
ACACTACCTTACANTTCNCAACTGNGCTTTGNTTTNTTTGNTTNGGTGTTGNGTAGTTATTTATTA

AAANG

SEQ ID NO: 86 ACGCGGGGAAAAATGGGGAGCAGGAGGCTGACAATGAGGTAGACGAAGAA  
GAGGAAGAAGGTGGGGAGGAAGAGGAGGAGGAAGAAGAAGGTGATGGTGAGGAAGAGGATGGA  
GATGAAGATGAGGAAGCTGANTCAGCTACGGGCAAGCGGGCAGCTGAAGATGATGAGGATGACG  
ATGTCGATACCAATAAGCAGANNACCGNCGAGGATGACTAGACAGCAAAAAAGGAAAAAGTTANA  
CTAAAAAAGGCGCCGCTGACCTATTCCCCTCCACTTCCCGTCTCATATCTTAACGTGGTCA  
CCTTCNANTAGAGANGCCCCCNCCCCGTGGCTGTCCCCCNNTTTANACNCGCTTTCNCCACCAC  
CCAANCTTGATATTTTCNAGGGGANGAANANACCAANTTTCAGGNCTTNTTTTTTAAANNNTT  
GNCNNANNCTNTGCTNATTTCTNCNTTGCNCNTTNTTTTGGTTCCNACTTNTCCANCTTGNNTTAT  
TGTTATTTTNTTTTTT

SEQ ID NO: 87 ACTCTGTATACACACATGAGAATGACAGTGACAAAGGCAAATAATGTCTTAG  
TATTAACCTGATCACTCAACATAATTTATTTTGACTTTTTTCTTTTACAGCAAGGGGAGGCTCC  
CTATTGCAACTGTTCTGACCCTATATTTTAGAGGTTTTGGGATATTAGGCTTATACACTAATAAAAT  
ACAGGGAAATATTTATTTTTCATAAAGTATTAGATGGTAGCTTTAGAAAGGGGTGTTTTTGTGCAG  
AGGGTAAGGCAGAGTAGGTAGTTGCTCTATTATGACTTTTCTTGGTTCAAGCAAAATAAAACCGC  
AAATGTTTATTTAAGNTGTTTTTGAAGAGAATCATAACCTGGATCAAACNTATTTCAAAGNNTT  
GGAGAAGNTTNTGTGCAAAGGAATTTNNAACNTAGNC

SEQ ID NO: 88 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGTTGTTTTGANACANAGTCTCA  
CTNTGTTGCTAGGCTGGAGTGCANTAGCATGATCTCGGCCACTGCAACCTCCGCCTCCNAAGTT  
NAAAGTATTCCCCTGCCTCAGCCTCCCGAATAGNGGGATTACAGGTGAGCATNATCATGCCAG  
CTGATTTTTGINTTTTAAAGTAAAAACAGGGTTNACCATATTGGCCAGGCTGGTCTTGAACCTCTG  
ACCTNAGGNGATCCGCCTGCTCGGCCTCCCAAAGGCTGGGATTACAGGCATGCNCCCCATGCCCG  
GTTATTTGTTGNAATANAAANACTTGGNGCCTCCTNTCCGAATCCANTNANNGAACAAGAAGG  
GGGCTCCATGCTCTNAATAANACCTTNTTCCCTCCCTATTTTAACTTTNCAGGGTACAAAAANGT  
TTATTNGGGCCATTTCAAANGGACCNGACCTCAANGAANGAATNCTTNATANAAAAACAATTGA  
TGTCCCCCTTCTTTCTTGGTTTTTTTAAAT

SEQ ID NO: 89 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNAGAANANATGGGGTTTNTCCATGTT  
GGTCAGGCTGGTCTNNAACTCCCTACCTNAGGTTATCCACCCGCCTTGGTATNNNAAAGTGCTGGG  
ATTAAGGCGTNANCCACANTGCCCGGCCTTTTTTTTTTTTAAAAAANAACNGACNAATATTT  
TACAAGGGAAACAAAATTNNAATTCCATGNNACNTNAGNATACCTNCACAGGGAGGAAAGGA  
CAAANTNTNTAGGATCTGGACCTGAAANCCAA

SEQ ID NO: 90 ACGCGGGAAAGGGACATTTCAAAGCCTATTGATGCTTGATAGTAAAAACCCTA  
ATATCCCATGATAAACACTAGAAACCAGCTCTCAGTGAAAAATGCTTTGCAATGTGTGGGTTTGTCC  
CACAGAGTTTACTTCTTTTGTATTCAANAATGTTGAAATACTTCTTTTTTGTAGAATCTATGTGG  
GGAAATTTGCGAGACCATTGAGGCCTATAGTAAAAAACTGAGTATCTTGCAATAAAAACTGGAAA  
CATGCGATCTGTGAAGATACATCTTCACAGATTCTGTGGATTCTACAGAGTAAAAGCTTT  
GTTTATTCAGCTGGTTGGAAACACTCTTGTGGAAGAATATATGAGCAGACATTTAAAGNCCCAAN  
GANGCCTTANNAAAAACAGGATTTCCTGAAAAAAGGNCANANTNCTNTNT  
GGANCNTCTGCTNTCCCATNNCNGNTAATGGTGNTTTTTGAAAACCTGGCTAANCTAA

SEQ ID NO: 91 GAGTCTAGCTCCCACCAGCAGTCGGGCCTGCCTCCTGCAAAAGAGGGGAAAG  
AGAAGAANGAGAAACCCANGANCAAAACAGCCCATCATATTGCCAAANACATGGAACGCTGGGC  
TAATATGTTTGAATAATCACAAAAGAAAACCTTTAAAAATAGNCTNTCATCCTGNCAATTCCTTTGA  
TGGAANAANAAAAGGANAAGAATCTNTTTGNA

SEQ ID NO: 92 CGTGGCGCGGCCGAGGTACTTTTGACCAAAAATTGACCAAAGTAAGAAAAAT  
GCAAGTTCTAAAAATAGACTAAGGATGCCTTTGCAGAACACCAAAGCTTCCCAAAGGAACTGGTA  
GGGAAAGTGGCCCCCTGTCTCCTGGAAGTGGAANAAGCCCTGCTCCCTGGCCTTTGGGTTTGCTT  
GGGGGCCACAATAAAATAANTNTTTGGCCCCCNNTTCCAGGGNCAAAAAA

SEQ ID NO: 93 ACTTTTTTTTTTTTTTTTTTTTTTTTTTCAANATGGAGTCTTGCTCTGTCACCTAGG  
CTGGANTTGGANTGCAATGGTGCAATCTGACTTACTGCAACCTTTGTCTCCCGGGTTNAAGCAATT  
CTCCTGCCTNAGTTCTGAGTAGCTGGGATTACAGGTGTGCACCACTCCTGGCATATTTTGT  
TTTTTGATAANACACTGTTTNACCATGTTGGCCAGGCTGGTCTNAACTCCTGACCTTAGGTGAT  
CCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCATGCCTGGCCAAAGTT  
TACTCTTTTACTTTTNTTTATTGGNGACATTTCTCCACGAAGTAGGAATTATCAAAAGTCTCATA

TCTGTATANATNAAAAAATGGAAACTGANACCCGNGTACCTTGGNCGGGACCACCC

SEQ ID NO: 94      ACTCCGAGGCTTTANATTNATTTTGGGTCTTTGGGGGGGACCTNTATCATTAC  
GCCTATNATCANTCCGAAAGTNATCATTTNTTGAACCTCANCNACCAACATCCCGTTGNCATANTC  
GGCTTTATAAGCGGCCCGCTTCCTTTGAATNNATNATCCTTGCAAANACCCATTGCCAANTGNTT  
TTTTGCCACTTGCNTNTCCGGGGAAATCATAACCNNTTATNATAGTCTTTGTTGCCTNGGNCATTNA  
TANNGAATCCCGGCCACCATTNAAACCATNTCCCGGGGGNGGGCCANTCCCGATACCCATCCCG  
AATCCNATCCC

SEQ ID NO: 95      GCCGTGGCGCGGCCGAGGTACTTTTGATATTAAAAGCTAATTTTAAAAACCC  
TTCTAATAAATTCATTACAGTTTAAACGAGCTTGGCCACACATGAAGGTNCCTTTTCAAAATTTAAA  
CCGGCTATACCTTCTAATNTCAATTTAAGTTGGGCTNACTCTTCCTNTTTCATGGNANAANATTT  
AATCCTTTAATTTAGGCCAAAAATAA

SEQ ID NO: 96      ACTTNTTTTTTTTTTTTTTTTTTTTNGGACAATTGTTATTTAGTTTTTATTTTCATAA  
TCATAAACTTAACTCTTGCAATCCAGCTAGGCATGGGGAGGGGAACAAGGGAAAAACATGGAACC  
CAAAGGGAACTGCAGCGAAAACCCAAAAATTTTTAGGAAACCTGNCAACAAATGGGGGGGNGGG  
GGTGCTCTCCTGAACCTCAAAAAGGATTAAGTGGNGGNTAAAAANAAACCCAAANTCAAACCTNTT  
CGAGTTGCCCCANGTCAACAATGGGGATCTTTTTGNTGGGCTTGCCATTCTGGACCCAAAAACCT  
CCATGGCTCACAATNTTCNNGCCTTTTTTCANTTTTCCAAA

SEQ ID NO: 97      ACAATCTTACTATCTTTCTTTTTCAGTTTGTGCCTTTAATTTCTCTGCATACTCC  
TGGTTCTTCTAGTTGCAATTCCAGAGAAATGAAATGTCTGACTTGATGTCTCAAGCATCAGATATA  
GCTTCTGGATGAAAATCTTGGCTTTGGAATCTGTCAAAAGATTAACCTTACTTTTTTCTGGCTCCC  
TGGTAGTTTTTCTGGGGGAAACCCAATTCTAAAAATTNCCTGAAACATGAAAATCTGGTTCAAAA  
NCCTTTTNTTCAAGGTAAAAACTCTTTGNTGGANNTANTTTCCGGTTTTTCTAACACTGGANACT  
GGGGAGCATGCAATTAAAAACTGCCTTNTCCTAGTANCCTTAAAAATGGNATCATGGGNGCTCAAA  
CCANAATAATTTGACTNTGCCTTAAAAAGTCTTNCCATCCGNAAATNGGTCAATCCCCTTTGGTAAG  
TGACCTGGAATTCGCTCTTNTCTTTTCTAACTTTAAGANTTCTAGGGATTAAACTTTTGCTTCCCA  
ATGGGCTGNGATTTAACAATANTTCNGGCATTAAANNCAATACCTGGCCCGGGCNGGCCGTTNAA  
AAGGGCCAAATTCNACCACACTTGGCG

SEQ ID NO: 98      ACTTNTTTTTTTTTTTTTTTTTTTTTTTNGCGGTGCCTCTAATACTGGGTGATGCTA  
GAGGTGATGTTTTTGGGTAAACAGGCGGGGGTAAGATTTGNCCGAGTTCCCTTTTACTTTTTTTTNA  
CCCTTTCCTTTATGAACCATGCCTTGTTGTTTGGGNTTGNCAAGTTGAGGGGAAATAATGACTTTGGTT  
GGGTTGAATTGAAAAATATTTGGGCCTGNTAAATTGCAANTCCAGTGNTTTNAATCCTGACNCCAG  
CTTATTGCGGANGGAAAAATGNTTTTCATNGTTACTTATACTAAANATTANTTTCTTNTCTNTNAGGG  
NGAATANATTGGNCCCAATTGGGGTGGGANGGAA

SEQ ID NO: 99      ACGTGTCTAAGTTCTAGAGCCTCCTGACGTGAGCATGGCTGAGAGTGAGGGA  
CCGCTCCCTGAAGGATCGTTCTGGGTAGGGGAAAACCTGGGAAAGTGGGAAAAAGTGCANAAGCG  
AAACACCATTCCTTTGGAAGAAGGAAAATCTTTTGATTCTTAAATTGGCTGGCCAAAGCTTGTTAC  
CAAAAACCTTGTAAAAAACCATTCCTCGGAATTGNANGGGAAAAANCTTTTTGTTTGGAAACACTT  
CANGGGTTTTTGAAACCAAGGGAAACCTGGANACCACCTTGAAGGAAATTAAGCGGTTNATTATN  
TTCTNCTGGCCAAGGCCCATGCTATTGGCCTAATTCTGCTGCTGGGCCCTTCACATAGGAGGAC  
ANAAAACCGGTGCATTGATCAAGGGTNTNTTTGGGAANTNANTCATGAACCATGGTCATTGGTC  
ANTCNCAAAAAAANTGGAAGGGCA

SEQ ID NO: 100      ACTACCTGGGGGGGTTTGCTTTCTGCCTTTTCTCTGGTTGGTCAACATCTTCT  
GGTTCTTCCGAGAGGCCTTCCTTGTCACCTACACAGAACAGAGCCCANATCAAAGGGCTATGTC  
TGGNCGCTCACTTGTNGGGCTTCCTTTTTGGGGTGGATGGGTGCTCACCTTCCTGGANTCACCATTT  
TTCAANANTTTACCGNCCCCGCTGGGGGTGCCCTTGGGNGACTANNCTCTCTNCACCATNACC  
CNNTGGGNACCCANTGAACAACCTTCTTCACATACTGGGGGCCCTTNTTATTCCNCCAANGANA  
ATCNTCT

SEQ ID NO: 101      ACGAGTGGTGGACAACAGTGCCCTGGGGAACAGCCCATACCATCGGGCTCCT  
CGCTGCATCCATGTCTATAAGAAAGAATGGAGTGGGCAAGGTGGGGCNAACCAGATACTACTGGC  
CTTCAAGGGACAGAAGAAAAAGGCGCTCATTTGGGGGGCACTGGCNTGCCTGCCCCCNATGAAC  
CCCAAATTCGACTCCACAACCGTGGTCTTATTTGAGACAACGGGAACCTTGTGGGGACNCNATT  
TAAAACNCCATTCCCNACCACCTTNCAGCCGAAAGNCNATATTCCAANGNGCTTGCCATTGN  
TAAAAACTTTNNNTGANTTGAACCCANGCCTTTTGTGCAAGACTTGTGAATGGAACAANTTNTGA  
AAACCACCTTTTGCTTANGGACTTNGGAACCACATTGCTGCTCCCTTTCACACTGGGTNACAATTT

ATATCCTGNGAAGAAAATAAATTNTNTTCATTCAAAAAAAAAAAAAAAAAAAAAAAAAAANTCCTG  
CCGGGCGGGCCNNTTTAAAANGGCNAATTTCAANA

SEQ ID NO: 102 ACACCCGCACGAGGAGCGGGGACGGCGGGCGCAGAAGTGGGCCACCATATC  
TGGAAACTACAGTCTATGCTTTGAAGCGCAAAAGGGAATAAACATTTAAAGACTCCCCCGGGGAC  
CTGGAGGATGGACTTTTCCATGGTGGCCGGACAGCAGCTTACAATGAAAAATCAGAGACTGGTGC  
TCTTGGAGAAAATAAGTTGGCAAATTTCCNNTTAACCACANTGACTTCAAAATTTTAAAAAATAA  
TGANCGTCAGCTGTGTGAAGTCCTCCAAATTAAGTTGGGNTGTTTCTCTACCTTGTCTTTTCAANTT  
AAGGAAGGCAACANCAATTTTTTGCAGGTGTNAAAAAATGCTGCTCCTAGGGATAAAGTTTAA  
NCTGGGAAAGATGACCTCACACAATGCTGTGTGTTNCTGNGGTNAATGCANCCAATGAAAAATCT  
TTTGCTTGGGGAAGGCCTGGCCCTGGCCCTGTATAAAGCTGGNGGATTTGAATTCAAAAANAN  
AGCAACNAGTTTGTGGCCANATTGGNAAAGNGTCACTTGTGAAANTANCTNTCCNGGANCAGGG  
ANGCTTNCCTTGAAAAAATAATTCATTCTNTTGGGCCTNNGNGGATGGAATGGGGATAAACAGG  
GGTT

SEQ ID NO: 103 TNGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATAT  
CTGCAGAATTCCCCCCTTGAANCGGCGCCGGGCAGGTACAGGAAGCATGGCTGGGGAGGCCTCAG  
GAAACTGGCAATCACGGCAGAAGGCGAAGGGGAAGCAGGCACCGTTTCTTCTTAATATTCCTTTT  
GAAATTCCTTATGGTGCACANGTAGCCGTAAAAAATAACTGCTTCACACTGACTTGTTCATNCCTTT  
TGGGGTGGGGGTANGGGGT

SEQ ID NO: 104 ACNCGGGCACTCACAGACATGACACACTCACAGACATGACACGCTCACAGAC  
ATGACACGCCCAGACATGATGAACTCACAGACGTGACACACAGACNTGACGCACTCACAGACATGAC  
ACGCTCACAGACATGATGCACTCACAGACAGGACACGCCCANATGCTACGCACTCACAGACGTGA  
TGCACTCACAGACAGGACATGTGTGGCTCCACTCAGCACCCATACTTAGTCACCTGTGCCAGNA  
GCACGCATGTCTACACAGATCACATTNACAGACACTGTGACACAAAGTTACACAGTCATGTGCAC  
ATGCNCACACACACTTGGTCTTNTGCTNGACTGCCTGNGCANGACACACAGAAGGCCAAAAA  
AAAAAAAAAAAAAAAAAGTNCCTTGGCCGAAC

SEQ ID NO: 105 ACTTNTNTTTTTNTTTTTTTCTTTTTTAAATTTGAACACATATTTNATAATTT  
GNGATAATGGTTTTGNGGCTAACTCNAAAAAANGAACGGCCCCAATCTTNAAAAGTCTACNCTGA  
AATATTTACAATCACNTATTTAACAGGTTTCTAAAANNATNACCATATTTGGGATTNGCTTTC

SEQ ID NO: 106 ACTTCTATGANAAGCGTATGGCCACAGAANTTGCTGCTGACTCTCTGGGTGA  
ANAATGGANGGGTTATGTGGTCCGANCTCANGGTGGGAACACTACAAACNNGGNTTCCCCATGAAA  
NAGGGTGNCTTGCCCATGGCCGAGTCTCCTGTTACTGATTAAGGGGNACT

SEQ ID NO: 107 ACAAACAATGNNTATTTGTTTGTAAGTGCCAGGTTTATATTTANNTAAACAT  
TAANATNTGCGTTNAAGCAGTAAGGCCNCAATNTTTANCTTGGCTGTGCNNGNTAAGTGACTTGTT  
TAATCTN

SEQ ID NO: 108 ACGCGGGGCCTCTTTTTCCGGCTGGAACCATGGAGGGTGTAGAAGAGAAGAN  
NAAGGAGGTTCTGCTGTGCCATAAACCCCTNNNGAAAAAGCAGATGGAATTTGCNANAGCTGAAA  
GATCAACACGCCTGAGANAGATGTTTG

SEQ ID NO: 109 GCGTTGGGCCCTCTATNGCATNAATCGAGCGGCCGCCAGTGTGATGGATATC  
TGCAGAATTGCCCCITAGCTNATCCGGCCGAGGTACAANACNCTACGGGAACAGNTTGCCTCCCT  
NCCAGCCTCAACCACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAANACTCCAGTTCTTANA  
GGCGCTGNAGTNTTTTTTTTTT

SEQ ID NO: 110 TGGGCCCTCTANAGCATGNTCGAGCGGCCGCCAGNGTGATGGATATCTGNNN  
AATTCGCCCTTACCNTTTGCGCGGCCGACGTACTCGNNATGACCCCAATACACAAAATTAACCCN  
NTAANAAAATTNATTNACCACTCACTNATTCGACCACCCCTCCCCNTCCAACATCTCC

SEQ ID NO: 111 ACCCTCCAGAAATTGGTGACTTTNCTTTTGTGACTGACAACACTTATACTAAG  
CACCAAATCAGACAGATGGAATGAAGATTCTTAAGAGCTTTAACTTTGGTCTGGGGTCGGCCT  
CTACCTTTGANNTTCCTTCGGAGAAGCTTCTTAANATTGGAGAAGGTTGGATGTCAACCAACANTA  
NTTNGGCCAATTACTTGANGGACCTANTTATGTNNGGANCCTTAAACAANGGGNCNTTTNCTTCTTT  
TAANATNGNACAAAGGACCTTTNNCTAACCNNTNAAAATTTNNGGNTANTGGGGGGAT

- SEQ ID NO: 112 ACTTTTGGGTTTTTTTTTTTTTTTGTATAATCTATTCATGGATCTCCACTTT  
GATTTAATTTTAAATGTTTAAATCTCCTTTACAAAAAGTATACGTTAATATACCCGTCCTCAAGGG  
GGAAGTGGCACTGTGACCTTNCATTNAGTTTTCTAGAGGATGNGATCNAATNTCTNTCTTCTCCTNAA  
TATNAAAAAGGAAATTNGCTTAACGACC
- SEQ ID NO: 113 ACTTGCCCTTCCCCAGAAAAGCGGGACTTGCTGCTAAGGGTGAAAGGACCA  
AGGCAGTTTGTCCCTGNTGGTCTGACCCCTTGAAAACGTGGGTGTATAAATNAGAAAAGCATTCCC  
TTCAATGATTAAANNCCAANNNGGAAGGCTNGCTTNCCAATTCCTTNAACAANCNCCGGAANTTTTT  
TTNGTTCNACGATTAAAANCNGTTTTNTTTTGGCCTTCCANAANATNAAATGGANTTTGTNATTA  
AATNANAANGTGACNTATTTTAAATTTAT
- SEQ ID NO: 114 ACGCGGGCAGTCAAGCTGGTTGCTCTGAAAGTAACCCAGCTTGTTGCTCTAA  
AATACCTCAGTAGCCTGAGTGTTATACTAGAGATCTAAAGGGGTTAACAGGGATAGGGGTGGGAA  
AGGTTAGAGACTCCTAGAAAATCTCTGGGTACCGTGATCTTCGGCCTCATTCTAATACCTGTCTTN  
GAACAGCTTTTTTCTNTTNGNGCTCTCTTNGCCTTAGCTACNTTTTTTATTATTGTATGCCTNCATNA  
CTTAATAAGNGATGGGAATGGGTTTTGANAAATCCGNAATTTATTTAAAAAGNTGGGTGGGNTTTT  
AAATACCTTTTTTTCAAATAAAAATTAAGCCNA
- SEQ ID NO: 115 ACTTTTTTTTTTTTTTTTTTTAAAGCAACCACATTTTTTAATCAATAAATATGAA  
CATTTGTTTTTAAACCAGTTACTATTAGGGCAGAAAAAACAGGCCNTTTTTCANAAATTTCAAAAN  
ATACNNCCTTTTGAAANNAANGGGCCAAAAATTTTGCCNTAAACTTTCCCTNNTTTNTGAACCT  
AAAGGTACCAANTTTTANCTAAAATAAGGGTTGGAATTATCNAAAAAGGTCCAGAATTTTTTTAA  
AGCTTTAACAGATTTTGCNATTAAGCCCAAACAGATTGGTTTANCCATTCTNANACTATNTGCCAC  
CCAATGGAATAAATGATTGGATGGGAAAATTGGTGAGCATTAGNGAATACCTATGGTCACTTATGG  
GCCCCGCTTACCTTCATTTCTTGGTCTTTTTCCGGCTTACTGNCCTTGNTNTTTCCAAGGCCTTGGN  
GANGGGGCTTCCAAAAAGCCTGGGCANTAANGGGCTGGTGGAATGTGANGGGCAGGATAGGG  
GAACCCANGGTTTTCTTTNTTTGNCTAACNTACTCCTTTTANCAAGGGGTTTGGTCTTGGCANNG  
GTTCTTTTCCCAATCCTTTGGGGGGCCTAACTCCCCTAGGGANCCCCCNCAATNGGTTCCAGCT  
CCCTTN
- SEQ ID NO: 116 ACTTCTTTGTGTAAAGTATTAGCCACTGTTTTTAGATCTAGTTAATAGGTTCT  
ATTTAATTTGCTCAACATTTACTGGAATGGGTGGAGTGAAAAAACTGATGCATACTGGGGAAAT  
ATCTACCATTTTTTAAAGATAATGGTTAATTAGGAAAAGANCCCTTTTAAGGAATTTATACCAGGG  
NTGAAATGGAATCCAGTTTTTTCAGAAATNCACCAACCTGGTGGATGGGTNTTAAAANGTGGCCTCC  
AAGGTAAGNATTCNGATTATATTAATGGATAACTGGGTGGTTTTANAATATGTAAGCCGGTTCT  
AATCTACNAGGTTACATATCGNCGTTNCTCTGNNATAATAGAAAGGTTAGGGGATGGATCACT  
TGGGGGGTTGTGGATAAAGGCCCTGNGTTGGTTCCCTCCTCTGGATNCCAATAGGGGGNATGCTA  
CTTTTTTNGGGGACCTNATTCCTTGAAGATTGGGCTTGCTATCAGGGGTTTATCCTTTTTCTCCA  
AAANNAANNAACCTCCTGCTGCTNCCATNTTAAAGCTNAAAATTTNNCTTTTGGCNGAANAATC  
NCCTAATTCAGGNANTCCAAATTTTTGGGGGNAAAANCNTCTGGGANAATNTATTGGACCNCATT  
NANGTTGGG
- SEQ ID NO: 117 ACTTTTTTTTTTTTTTTTTTNTTTTTTTTNGCATCAAAAAGCTTTATTTCCATT  
TGGTCCAAGGCTTGTTAGGATAGTTAAAAAAGCTGCCTATTGGCTGGAGGGAGAGGCTTAGGCAA  
AANCCCTATTACTTTGCAAGGGGCCCTTAAAAGTCGNTGGGCTCAAAAGGNCTTTAATNCTCCTT  
TAAAANTTANGCCTTTCGNAANAAATTCTNNGCCAANCAANGCTTNGGGCCACNCAA
- SEQ ID NO: 118 ACCACTTGAAGCCAGAATAGTTNGNTTATGTGGAAACCACGGGACCNGGAAA  
ATTTTCATCTTNATNGAAGATTTCGANGGTTTGGANATTTAAATTTAAGTNACCCANATNCCATNNA  
GCTACGGTCTTTATCCNCAGAGCCGTGGCTAAAAATAAANAC
- SEQ ID NO: 119 ACACCTGATTTCAGATTCCACCTGGGATTTCGACAAATTTTTCTTTTTTTGTTTTT  
CTATTTTCTTTGGGATCTTCTCTTGTTCGTCTTTAAATCTCTTATTTACCCCTTTTAACTTTCTTT  
AATTCCTTTGTTTCTTTTACTTCATCTTTTTTGGGCTTTTTGGACTCCTTTAGTTCCTTTGGCTTCA  
GAAATTCCTTTCTTTGTCTTAAATCAAAAACCTAACTTTCCAAGGAGCTCTCCAGGTCTGGTTTGG  
ACTTGTCTTTAGCTNCCCCGGCCTTGGCTTTTTCTTTTTTCAGATCATCTGGGGCTTTTCTTCTCTC  
TGCCTCAAATTTTCTTTTTCTTTTNGGGGAAGATCTTCTTTTGTCTCACTTTGCTGATCCGCTATC  
ANAGTTCGCCTAAAATATGTCGTATTTAAGGATANGCCTCTGAATATCCTTCTGACTGCTTTGGC  
TTTGTCTCTGCAATTTCTCCCTAAATTCANAAGCNCTTCTTTTACAAGTCTCCAGGGNGAATCTC  
GGGCTCCAAGGCCCGCGTCTGGGCCCCGACCACNCTTAAGGG

SEQ ID NO: 120 GTACCTCAAGGTTCTCAGGACCTCCTTTCCCCAGATCTTAGGGTCCTGCCCTG  
TGGGTCTCCTGTGTCCAGGGGAGAGGATCTGGGGAGTAAAAATTGTGAAGTGGCAAATCCCTGTA  
CCCGGTTTTNTTGGATTTTCCAATGAAGGGCTGTCTCACATCCCACCTNCCACTAGGGTNAATGAG  
AGGA

SEQ ID NO: 121 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAATTTTTTTTTTTTTTTTTTAAAC  
CTTCTTCCAGTTTCTTGATTTTTTAAGGNNTTNAATAACATTTATTTTAANATTTTCTATNTATAC  
ATTTGCTACGGAAAATTGAAATCCGCNCACTGAAATATCCTTTTATTGCAACTNAAATTTTCAAAA  
ACCAAAAAATCAAATTATTTATGCTCTANCCAAATTATGNAANGGTTTCTTTTTCTTTTNTNTGTCC  
CCAAGCTGGANTGTANTGGAACAATCANANTTCACTGGANCATNNGCCTCCNNGGCTCAANNGAN  
NTTANTTGCTTAACCCCCCGAGTNGCTTGGAANTANCNGGTGAACAAAACCACANCCAAGNTTA  
TTTTTTTATGTTTTTGNTAAAACCGGGGCTTTGCT

SEQ ID NO: 122 ACGCGGGTGCCATTCCCTCCTCTTCTGGATTTTTTTTCTTTGACCATATCAAGC  
TGAAGAGATGAACTTGTCTTCTCAAACCTTTGCATCAAATTAAGAGTAAAAACAATAAAAGTAAT  
ATCAAATGAGGCAGGCCAACTAAAACAAGAAATAGGTTAGAGAATATTGACTCTCCGCCGGGCAT  
GGTGGCTCATGCCTGTAATCCCAACAGTTTAGGAGGCTGAGATGGGCAGATCACTTGAGGNCAGG  
AGTTCAAGACCATGCCTGGCCAAACANTGGTTGAAACCCTTGTCTCTACTAAAAAATACAAAAATT  
AGCCAGGCGTGGTANTGGGCCCTGNAATCCNAGCTACTTNGGANGCTGANGCAGGA

SEQ ID NO: 123 GTACTCTNTTTATACTNTAATCTGGNGGATANCTATTTTAATTTATGTTATTC  
AATCCATATCTTCNGTTGNTTCNCAAAAACTGANTTTACTACNANGTATATATTTAATAAACNT  
ATTAATTATAANTTTTNGNATTTAA

SEQ ID NO: 124 ACTTTTTTTTTTTTTTTTTTTTTNTGGGGAAAAATCCTTTTCTTTACAAACTTCCAT  
CAGTTTAGGAGTCAGTCTGTATGCCTTTAGTGAGAGAGATCCTTGGGCAAGTTTTATTGGGATCA  
TAAATGANAAACGACAGATTCTTCAATGGGCNTGCTGGTNACTAAACTGGAGANTCCNNGAGTG  
GCCCCGNTTACNAATGAGCCATANTATGNNGGACTGAATACCNACCCACGTAAGANNATCA  
ATGTTTANNTTGGCGNAANGCTCCCNATTATTTCCATCTNAATTTGGNNCTNATTCAAAGTNCAC  
NNTANTCC

SEQ ID NO: 125 ACTTTTTTTTTTTTTTTTTTTTTGTATTNNGGCTNCCTTAAACAGTTGGACCTTC  
CATTTTAATCCAATTTGTAAATCTTTGATTCTAATTTGATGCTCTGACCACTTGCATTTTATGTAAT  
CAATTGAATATGGGTTNGGNCTTAAACGTTACTATTTGGTTTNNTTGAATCCGTGTTTTNTTTCNAN  
CATTTTTCCTTTTTTTAAAGCCTTATTTTGAATTTGNNTNTTCTTNNAATTCNCTTTNANCCTCNCN  
NAGGGTTNGGTACCAAAAAT

SEQ ID NO: 126 ACCCTTGCCTTTCTCACATCATNAGATCAAGTCACTCTTTGTGCATCCCTTCCT  
GGCTGAGCGCATCATTTCCATGTTGAACTACTTCCTGCAACACCTGGTTGGCCCCAAGATGGGNGC  
CTAAAAAGTANAANGACTTCAGCATATNTGANNTCAACCNAATCTGTANGTATCAGATATCTGNA  
CTTA

SEQ ID NO: 127 ACAGTGTGGCTCATGCCTGTAATCCCAGCACTTCGGGAGGCTGAGGTGGGAC  
AATTACTTGAGTCCAGGAGTTTGAGACCAGGTTGGGCAATGTGATGAAACCCTGTCTCTACAAAA  
AATACAAGAATTAGCTGAGTGTGGTGGCAGATGCCTGTANTAGCCACAGCTACTTGGGAGGCTAA  
GGAGGGAGAACCCTTGAGCCCAGGAGGTCAAGGCTGTAGTGTGCTGTGATCGCGCCANTGCATT  
CCAGCCTGNGTGACAGAGCAAGACCCTGTCTTAAATACANTNAAAAATNAAAAAANATNNTT

SEQ ID NO: 128 ACACGGCAGTCTTAGAGAAGCAAATGGCTCAGATGATGATAATTAAGAGTAG  
CCAACATTAAAGTTAATTTTTAAAAATACAGTTAGGTGTTTATATTATTAGTTATTATAATTCTGC  
AATCCTCTTGCTCAGGAAGTGATACAACCTTTTTTAAAAATTATACTTGTTATTATTCTACTTCTCT  
CTTCTATGACAACTCTAGTGCAATATTAGAGTTTCATTTATTCCACAATATATTTTACTGTTTTTTT  
CTATTTATCTGGCTTACTAGTAAGTCTTTTAGATTAACATCCAGTTCATTTTCTGGAAATATACC  
TGGAATTTAGTCCATTTCTAATGAAAAACAGCTCAAATTTTAGGGTAACTAATTTGTAAAGAGTT  
AACATTATATTTGGATCCAATTTTCTGANATATTGAAATAGTGAAATCCCAAATAAATGATNTGA  
ATTATTTTACTTTCAAATGTATANTTAATTGCTAACATGTTTTTATTTTCATATATTGNTG

SEQ ID NO: 129 ACAGTAGAAACAAGCAGAGCTACTGATACTCTCACAGCTCATTTTCAGTTTGTG  
TTCCTTATCTGTATGAAAGGGGACCATAGAGAGAGGTTGAATTAGTTTCAATACAGCCCTAAGCAC  
TTCTTATGGTGTTTTTTGAATTACTGCTCAACAGTTCAGTCCAGTTAAAGTAAATAAGCAAATTAA  
ACGTAAGAGAATGAACTTGGCAACACCTTAGATTATAATTCTGATTTTAAACAAATTCATTCTATT

AATGCACATAATATTGATACTGCGTTGTAAACATTTAGGCCCGTAAAATTACTGAAGGACTGTAGA  
 ATGAAAGAGAATCNCANAATAAACTTAAGGTTANAGAACACTTAATGTTTCCTGCCTNAAGTGNG  
 ATGCTATACTTGCAAAATNTTATTGNGAGGNAGAAATATTATNNTTTNAGCATATAACCATTGTTN  
 TCATT

SEQ ID NO: 130 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGANANACAGGGTCTTGCCACTTTGCC  
 AGGTTGGTCTCAAACCTCTGGCTCAANANATTTCTGCCTTGTGAATCACTGANCAAACCCAGC  
 ATGCATTCTTCTAATTTTTTTTAATTTATTTTTCTTATTTTTTAAGTTTAAAGGAATGTTAAATTCC  
 ANANATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGAAATGGAGTNTTACTNTNTTGCCAGGCT  
 GGAATGCAGNGGCAAAATCTCGGCTCACTGCAATNTTGCCTCCCGGGTTNAAAACATTNTCCTGC  
 CTAAGCCTCTGANTAGCTGGGATTACAGGNGNGCNCACCACACCCAGCGAATTTTTGTNTTTTT  
 TTTAGTAAAGTCAGGGTTTCCCATNTTGACCAGGCTGGTCTCAAACCTCTGACCTTNGATNTGC  
 CCCCCCTTATTNTCCCAAAGTCTGGGATTACAGGCGTNAGCCACCGNCCTGGCCAATATTTNCCGC  
 NCCCTNCCGGGCGGNCGTTTAAAGGGCAANTTCCANCACACTGGGNGGNCNTTNTNGTGGATN  
 CNANCTNGGTNCCAACTTGNGGTAANNATGGGGNATANCTNGTTCCTGTGNGAAATGTTTCCC  
 C

SEQ ID NO: 131 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAATGCAACAACCTTATTGAAAGGAAAGTG  
 CAATGAAATTTGTTGAAACCTTAAAGGGGAAACITAAACACCCCNCTCANGCNCANGACCAAG  
 TGCAAAATGGACTCTTNTGGATGTTGTANNANACNTGGTGCANTCNTTTTCTAATCTGATTCCTA  
 AATTAATCAACCTNCTGATNAAGAGGGATNCNTTCATATATTTANNATNTTGTNTTGANT  
 TTTAAATGGTNATNATT

SEQ ID NO: 132 ACCTACATCAGATCTAACCTTGATCCCAGCAATGTGGATTCCCTCTTCTACGC  
 TGCCAGGCCAGCCAGGCCCTCTCAGGATGTAGAGATCTCTATTNCAAATGAGACCAANGATCTG  
 CTTCTGGCAGCTNGTCANTGAANGACTCATNCTGTTNCCNGATCTACCATCCANTANNNGCTNTT  
 AAGGGGCTTNGGCCTTNNNANCTATCTCANTAANCNCTCATGTGNCNTANTCTGATCATGGTTAT  
 CTANGA

SEQ ID NO: 133 ACAAATGTTTTTTTATTCAAANGTNCAAAATAAATTATCTGTAGGCATGGACA  
 ATGACAGCAGTAAACCATTATATATTNGTCAACTGAAACCAGTNACTGATGGTTATAGTGATTN  
 AGCCGCCTTTTCTTTATTTTNTCCAAGTACTTCTCTGAAAGTTATTGGTGAAGGAACACTGCCTT  
 GGCCTTCTGCANANNTCATTAAANTAAAGNAAANCCCTAGCNANGAGNTTAAACATGCCACNTC  
 CATTACAACTCCATTTNACCANTGGNNCCATTCCAAGGGCCCTNTNTNTTAAAGAATTTNTNG  
 NNAC

SEQ ID NO: 134 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGGATTTANTAAAATAAATGTAT  
 TTTTAAATNTTTTTAGGAAGCTGCAANACTGGCAACGTGATTGCTTCTTGATAAAGTCTGCAATG  
 TTAAGTTCTATATAGGCTCCCTTCTGTGCTTCTGTTTTGCAAAAACCTTGATTAAATTTATTTCCCT  
 CAGCACCTGATCCTTGCCCTTGACCCACTGGCCATTTTGGATNGGACCACGTGATACANGTTGGA  
 CCAAGTTGGTTGCATTAATCNATTAATCTCTCTGAATCTGGCAATCANCAACCATCANTCTTTTTN  
 TTTTTTACAAAATAACCACCCATTTCNCCAAATTTCTTNTTGAANCGTTAAAAGGCATATTTTCC  
 ACCGTTTATNCTATTTCTTAATAGGCCTTCTTGATGAAACCTTGAAATAATCTGNGGAATCNTG  
 NCTCAAACTNNGGGCAAAAACAAAAAATTGAGGGCANTNAACATGCCCTGNGCNTGNTTATACATA  
 TNCCTAAAAAATTTCTCCNTAAAAAANANTNGGNGGCCAAAATTTTNCNNCCTTTTTTTGGGAA  
 ATTAAATTCNGTTATAT

SEQ ID NO: 135 ACATAGTAAACTGTGGGTATTCAGGGAGATAAAAGTTTTTTGTTTGTGTTT  
 GTTGTGTTTTAAATGAGGAGTCAATAATTGTTTTTCAGATAATTATCCTTGGCTACAAAGATGAATA  
 ACAGGGGTAATACCACTCTGAGGTTGGACAGTTGTTGGGCAGTTGTTGGGCAGACGTCCTCACAG  
 AAGTATTGTGTGCCTATATGTGTGCGTATAAAGTTGCAATGGCCTTTTTGCAAGGTTGTGGTTTTTG  
 CACCTTTTGTGCATGAGAGCCCTCCCTTCTGTCTTCTCAGTTCTACTTGTGAGGGTTTTTGCCAC  
 AACTCCATTTTGATTCTGACAACTTTCAGAAAGCCTTAGGAACTAGACAGTGGAATTATGGGATTT  
 GCTCTTTAGCCAGTTGCATTTTATTTGAACTGGCTTTAGCAATGAGCACTAATGAAAATTAGCCAC  
 TGCTGCTGGGTGCAGTGGCTCATGCCTATGATCCACACTTTGGATCACCTGAGTCAGGATTTCAA  
 CCACCTGCCACATGGNAAACCCNCTCTACTAAAAATAAAAAATTACCCGGCATGGGNGGTGGAC  
 CCTATAATAGCCACTACTGG

SEQ ID NO: 136 ACCTTTTTTTTTTTTTTTTTTTTTTTTNGGTAAACANGGCGGGGTAAAGATTTGCCGA  
 GTTCCTTTTACTTTTTTAACTTTTCCTTATNAGCATGCCTGTGTTGGGTTGACAGTGANGGTAATA  
 ATGACTTGTGTTGATTGGAAATATNGGGCTGNAATTGNCAATCCAGTGTTTATAATCTGACCC  
 CANNCTTATTCGNAGGANAATTCTNNNATTGNTACTTNTACACAAACACAATTNANNTTATNAG

CATCATAAAATCGNCCCA

SEQ ID NO: 137 ACTTGACCCACAGCCGTCNNGGATGAGCCGCTTCTCAGCCACCATGTCTTCA  
AATTCATCATCTGNACNTGGTGAANCCCCACTTCTTNGNNAANNGGATCNTNTTNGCGGCTAT  
TGAANACTTGTC

SEQ ID NO: 138 ACACAGCTGTCAGGGAAAAGTCCTGATGGCCACAGTGAAAAANGTCATGGTTN  
GACAAAANATTCATNGCANCTTAGCATGGNTCANACCAGTNTCNTACNTAAATNNCTAGNTCNCA  
ATCTAGTTTTINACCACATCTGCTTCAAAAATAGCACACCAAACCTCCCTTCCAAAGTGCATCGTTAC  
ACTTGACCATCGTGGAAGAAATGGAAGAACAAGGATGGATTTTGGCTGGCTGGAAGTCACATCT  
TGGGGAAGCTGGCCANGTTTGGCATTCCACAAGGCNTTGTCTTATTTTCNAGCCATNAGGANNTA  
TTCCTTNGTTCCCAAAGTTTCTCCAGNTTNTTTTAAATTATCANGGCTTGNTTTCCTT

SEQ ID NO: 139 ACGCGGGGGCGAAGGCGGGGTGCGCGCTGCCGGGTGAAATCGTAGGACAGT  
GAAGATGCTGCTGGAATTGTCCGAGGAGCATAAGGAACACCTGGCCTTCTGCCTCAANTGGACA  
GCGCGGTGGCTCCAANTTTGGGCGGATTGCNGTGANATTTNTGAACTCTGGCNNAACNCAAA  
ANTNTANTGAAGGNNCNTCCTGAAAATC

SEQ ID NO: 140 ACTTTTTTTTTTTTTTTTTTTTTTCCGGTTTTTTTTTTTTTTTTTTTTTAAA  
AAAGGAAAACCCGGTANTGATNTCGGGGTGAGGGATAGGAGGAAAATGGGGGATAGGNNTTN  
AACNTGAGGGNGTTTTCTCGNGTNAATGAGGGTTTTATGTTGTAAATGNGGGGGGTGAGGGACCC  
CCNTTGTNTGTGGTNANTTTTTTAAGGNAATNTTGGGTTGANNATANTTTGTNGGNTCCTTTATA

SEQ ID NO: 141 ACTTTTTTTTTTTTTTTTTTTTTTGGANAAGGAANAGGTTTTTATTCGGCCC  
GGAGCCNTCNGCNACTCCCGTCTNAAGAGCCNANCTCCCNAAAAANAATTCCTANCCCTTTT  
AAGAGCTTACAACTTAAGGGGTCCACGTGAAAGGGTNATANTAGATCAAGTAANCGTGAGGAAC  
NTNANTGGGGGCTACACACNTTGGCCTTTTGGACAAAAANTTTNACGGTNTTCTGACACANTGT  
CTTGGGTTT

SEQ ID NO: 142 ACTTTTTTTTTTTTTTTTTTTTTTGGCTGGATTGCTTTTATAGGANAG  
TGAAGGGAAACCTCATGTTTCTAGGATCAANCTGNGGTTTNAANATTCATCCAANTATTTCTCAC  
ATATTCCTNTTCAAAAATCAAGTTTCCAAAACCTCTGCTCTGTTTCTCTGTTGACCACACAAAA  
CCTTCCAGTANTTGAATAAAAAAAGGCACNTTNTCACTGGGAAAAAAAAGGCACCCAGGCCAA  
AAATTACCTTTTCAAGATTNTGNTGCCAATGATTAACACANAAGGAATNTGCCNTTGGTTNTAT  
NCTGCTGCGGNAACCTAATNCTTGNNTTGNGTNANATTTGNATCCAATTGGNGATTTANGATGAGG  
GAAAGCNGGATTTTNTCCNACTCCCNNTTANCTCNTAAGNNAGCTGGTNAANTGGGATNTTTGN  
ANGTNTGTATCAAANCTTCTTGANNNANATTGTTNNAAATTTAATTACNANNCCANTAAAAAAA  
AATACTT

SEQ ID NO: 143 ACCGAGTGTGGCACCTAGGACAGCAGGCAGTAGTGCAGATAAGGTGTGACTC  
TTTCTAGCATAGCCAGGGGGCATGGCTACCCTCATATATCCCCAGGCCTTCCCTAGACTCTAATGG  
TTCCAAAAGCCACAGAAGCAGTTTATGTAGTTTGTAAACACAGTATCTGACCTTCCTAATTANA  
CCAAATGTTTAAATTTTGAAGACATTTTATTTTACCAATAATCTTTAAAACTATCTTTATCTCCA  
AAGATTAAAGTCATGGGAATAAAAGGCATFAAAATTTCTACTTTCCTGAGAAATATTTAAGTGCT  
TATTTTCTTTAAACCAATTAATTAGAGATCTTTTATATAAACATTATATACACAACACATATAAAT  
ATGCACACAGATACACAGATAGAAGAGCTTTTTTTTTTTTTTTTTTTTAAAGACANAATCTCACT  
CCGTCCCTAGGCTGGAGNGGGAGTGCCATGATCTTGGCTCATACAGCCTNAACCTTCCANGTTCA  
AGCAATTTCTTGNNTTNAACCTCCCGANTAGCTTGGGACTACAGGCCCCCGCCCCACGTTCAAGN  
TAATTTTTTTTATTTTATAGNAAAAATGGGGGTTTTGC

SEQ ID NO: 144 CTGAAGNAACTANCNTCAANAAGTAGCCTCTGTATGGGAATAGAGCTAAGGA  
GGATGCTAAGGCTCAGGCGAACTGACCCCTTGACAGCCTAACATGGAGGCTTTTATCTTTGAATTT  
AACACATTTGTCAACTTGACAGGGAGCTTGGGCTGCAGATTCTGCCCCTGTGAGACTCTGAGGCC  
CGGCAGAAAGAGCCCAGGCATGGGAGTCAGACTCATGGGAGGGTGTGGGGGTAAATCCTGGCCA  
TGCAAACTTCTGCACTAGGATCTTTACTTCCCTGGCTAGAATNCTATNACTATAAAAANTGGGGG  
GAATAAATGACTGGNTCANCGGGGTTTTTAAAGAAAAATTNAAANGGGTTAAATTTATGCAANAA  
ACAGGGATTGTGNGAATNAAACNGANGGAATTCAAGCCAAAACCATGGGGGGAAAAAAAACC  
CCCTT

SEQ ID NO: 145 ACTGTTCAATAAAATTTAATTCATATAAATCAATTTTTTAAATTAATTATA  
AAAATATTAAGCATTTTTTAATTCAACATATATATATACTATGCTTAAAAAGTCTTGATCATGG



AGGTGATATTCACCATATTATATCTACTAATGACACAATGAACTTTTTTGAGGGATACATTTTTTAA  
 AATGATTCTGTAAATGACTAAAGTCTTTGGTAATAAAAGCTGGAGATTGGAGGTTACTTTTGAAAA  
 TAAAGAATTTGAAGGAGAGAATGTAGAAATGTGAAGCAGAACTTAGCCATGAGATGAAGACTA  
 AATTGAGAACGGAATGTTTCATATAAGGTAAAAAAATATATATACACCGTTTAATAATTGTTTAG  
 AGGGGGAAGGGAGGGCAGGGAGAGAAGGAGACAAGGGTTGAAAAACCNGCTGNTGAGTAGCAT  
 GCTCACTCCTGGGGGNNANATCATTTGTCCNCGGCCGNAGCCANGCTAAGGGCAAATTTCC

SEQ ID NO: 146 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTGGCTTTTAAAAATAACCTTTATTTTTT  
 AAAANTTANTNTGTGCNTTATAGGAAATNGAAAAACACAANCAAAAAACAAAGTCATTACAAATC  
 ACAANNATTTNANANTTTGACATATTCTTCNAAATCCTGTGTGTAGGCACAACATCCAATTTTATG  
 GGA CTGAAACTG

SEQ ID NO: 147 ACTTTTTTTTTTTTTTTTTTTTTTAAATATNACAAAATNAAAACTTANACACTTTAG  
 GGTAGCNGCTTAATACTTATCTTNGAAATCTATTGCTNATGCTAGGTCTAANGAGCAATGACTCA  
 ACCAGAAAAANTANTAAAGGCTGCCTTTNCCCTTTTNAAGTCTTATTAGCTTTATATACAAAATA  
 CAATGNTNTNACTAATACATTANTACTNAAAGGTGCTTAAAAAT

SEQ ID NO: 148 ACTTCCGGTGCTAAGGGNTNTCCGATTTGTAGAAGGCACAAATATTAATAGG  
 ATCACTTTTAGCTCTTGGAATGTCATCAATGCCTTAGCANATTCAAAGAGAANGAATCANCATAT  
 CCCTTACAGGAAATAGTAAGCTTACTCGCTTGTTAAAGGATTCTCTTGAGGAACTGTCAAATA  
 TAATNATAGCTGCTGTTANTCCTTNCCTGTATTCTACTATNACACATNTAACACTCTTAATTNGCT  
 AACCNNGNCAAAAGNANATTAAANTCTCTTTGAAATTGNA

SEQ ID NO: 149 ACGCGGGGCCATACCAGCCTAGGTGTGGAGCAAGAGGTAGGGAGGCCCTCG  
 TGGATATACACAAACACCCCANATACAAAATGGAGCATTGTGGTAGTGGTTAGGGTGTTTTATGN  
 AAACANITTAATTANATANTTCTATTCACTGA

SEQ ID NO: 150 GTACGCNNGGACGACGAAGATGATGAANATGATGATGATGAAGATGATGAG  
 GAGGAGGAAGAAGAGGAGGAGGAAGAGGTGGGGTGGGACGACAGTGAAATCTAGAGTAAAACC  
 AAGCTGGCCCAAGGTGTCCTGCAGGCTGTAATGCAGTTTAATCANAGTGCCATTTTTTTTTTTGTTT  
 AAATGATTTTAATTATTGNAATGCACAAATTTTTTAATNTGCAAATAAAAAGTTTAAAAACTTANA  
 ACNCAAAAATTNAAANAATTNTAAANTTTNAATNTNCTTGCTGGGCTGGNNATTNGAAAGGANA  
 TTTCCATCANNCTGCTTTNTTTNTTTTGTATCTNAACTCNGCNNTATTTGNTNNTGATNTTNTCTT  
 AG

SEQ ID NO: 151 ACTTTTTTTTTNTTANTTTGTTTTTGACATANATNTANTCTNTGGTNANGGTGG  
 CTGGAATATACCTGACCCACCATTTTNNANAANGACCCATNTNANGTCTGACCAATTGGGAGCAAAG  
 CCATGTTNACACTGACCTAATGCAAANTATGGAACCAATTGGGCTGGTTATACATTTCTGTTTCTTA  
 AATTATNNTCCAACNTGA

SEQ ID NO: 152 ACGCGGGAGACTGAAAAACTGCCTCATGCATGTGTTCTATTTATTGATATATG  
 CACATATGGCTGTGATTCACGTAAATTCATTTTTAAATGTTACTGAATTCACAGTATCTTTTCTCAA  
 CCATGTATTTTCTCAATAAATGAATACTAAGATGCANTTTTGAAAGTATAAAAAATAAGGAGCTGCT  
 AGAAAAACTAAGTTCTGCTTTTTTGTTTTTTTTATCTGTAAANTGGAGAAAAAATTNATAATTNNT  
 NAATATGGTTNNAACANCTTATTGCTTTTGTATCA

SEQ ID NO: 153 ACTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGNCTCACTCTATTGCTGAGGCT  
 GGAGTGCAATTGGCNCNATCTTGCTCACTGCAACTTCTNCCGCTGGGTTTAAGCGATTCTNCTGC  
 CTC AACCT

SEQ ID NO: 154 ACGCGGGGAGAAAGGAACACAGTAAACTGAATTGATCCGTTTAGAAGTTTAC  
 AATGAAGTTTCTTCTAATACTGCTCCTGCAGGCCACTGCTTTGGAGCTCTTNCNTGAACAGCTNT  
 ACATGCCTGGAAAATAATAATTGTCTATTNGGTGAAAGATACTTANAANAATTTNTTGCCTTTNN  
 TANANTTCATTNTATNAANAGANATNTTTNAATATTNGTATATNTT

SEQ ID NO: 155 ACTTATGTCCATTTCAAGTTTCCCCACCTATAAACAAGAGCCAATTTCTCTTATT  
 TCCCTGCTCTCCCCAGGTTGAAAAGGTGCTGGCCCCCTTGGAAAGATTGTATTGACTGTGTTGGGGA  
 TCTGGTGCCACCTGNTGNATGCCACAAGAAAGGCCTCTCCTGACTCCCAAGTTGTAACCCGTTTCC  
 ACCAAATCGACTTCCAAATAATATTATCAAATCATCATCTGTGCTTTTCTCTTGTTCAGACCAC  
 TTTTAAGGTGGAAAAAGGCAAAGAAGGCTTATATGTATTTCTTCCATAATGAGTCCNTCANAAAA

AAGTTNCTTCGGTGAAATTNTTGACCACNTTATGTTTNGGGGGACTCCCTATNGGATCA

SEQ ID NO: 156 ACCACTGNATTGATTAGNGGTGTATNTAAACANGGCTCCCTTCATTGCATCTG  
AGGACTTGTTTTCTTTTCTTTATTTTAAATCCTCTTAGTTTTAAATATATTGCCTAGAGACTCATT  
ACTACCCAGTTTGGGTTTTTTGGGANAAATGTAACCTGNACAGTTACCTTTTCAATTA AAAAAGACA  
CTTACCCCAAAAANAAAAAATTA AAAAANAANTNCCGTCNCNTTTTTTTGTGT

SEQ ID NO: 157 ACCTGGAGGCTCAACGGTTTAAGCTTCACCACAAAAGCNAAATGGGCACACC  
ACAGGGAGAAAACCTGGTTGTCCTGGATGTTTCGATAANTTGGTCGTNGTCTTGGNNTTNATNTNTTC  
TNCTNAT

SEQ ID NO: 158 ACTGTAAAAGTTCTGACACAAGACAGTGGCAGTGGTTACTTTTCATCGACTTT  
AGCATGTGATCTCAGGGACTCAAGACATACGCTAAGTTCTATTCTGAGTTTTGGGCAACAGAAGC  
AGTGACAGATATTTTCTGAATGAANAAATTTTAAAGGTGTTTTCAAGCATTTTAAAAAGTNTTGGCCA  
ACACACTNTTNGTTGTTGCTCACAAGTCACNTTTGNTGCCCCAAAGAATTTAAAGAACCCTTTAAANG  
TCTACAACNTNTTACTTTACCAAAGAANTAACNTNTTANTNGAAGGGGTTAANTTTTTNTTTNATT  
AAAAAATTCACACTTANCTTTTT

SEQ ID NO: 159 ACAACTATGATACATAAATTA AAAAATACAAAAAAGGAGGGGGCAGGC  
ATGGTGGCTTATGCCTGTAATCCCAGCACTTTGGGAGGCCAANGCANGGTGATTGCTTGAACCCA  
GGAGTTGAGACCACCTGGCANCATATTAANACTCCTTCTCTACAAAANTTTTTAAAAATAATNTC  
CAGGTNTGGGNGGNTCAACTCNTTTTANCCCTANNTATTTTNTNATACCTTTATTCTNNTAATTTCTN  
TTAANGNTTTNTATNATTAATAATC

SEQ ID NO: 160 ACCCCCTTCTCCACGTAGCCACGGCTTCCCCTACTATCAACATCCTGCACTA  
GAATGGACATTTGTTACAATTATGAACCTACATTGACACATCATTATCACCCAAAGTCTGTAGTTA  
ACTGTTAGGGTTCCTCTTGGGGGTTGGNAAGTTCTGGGGGCTTNGATAAANGCGTAAGNGGTTTC  
CNGTTTTTAAGGGCTCAAAAANGTCTTGCTTATNAGNGATTATTGCTNGTAGGAATCTATGAGGCTA  
NAACTTATCTNACTGCCCCAGCATAAATGTCTTATACCATATCAGCTTTAAATAACCCATTTTAA  
GTATCATTGANAATAAGACCTTTTTTTTTTTTTTTTTTTTGGCACTAAAATTTTATGNTTTTTTAA  
NCCTTAGCCGNTTGANAANNGAANAACCTGGAATGACAGCGGGNGGGANGAACNGAANN  
AANGATNNNATNGGGGGGNCAGGTGGGGGGAATTAACNNAANTTTNTTTTTTCANGGGGGAGTC  
CTNACCTGGCNTNATGCCATGAATGAGTNGCACAC

SEQ ID NO: 161 ACGCGGGAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCA  
AAACATCATCTTTTATCACCTGACCCAGGAAGTTTTTCATTGGGAAAAGGGATTTGAACCTGGTGG  
TTACTAACCATTTTAAAAGACCACCACCAAGGGAAACCAAATCTTTTCTGAAAGAAAGTAAAT  
GAATACCTTTTGGGTNAATGGAATTGAAATNCAAAAANAATTCTGNNCTTTATTGACAACNAATGG  
NTGTAATTNATTGTTGTAGATAAACCTCCTNTTATNCANCNAAACCCCTGTTGGGAAATGAATC  
AACTNCTTGAAATACTTTATTAANTAATCCAATTNCTTCCNAANTTTAAGTTTGTCTGTTACNCCC

SEQ ID NO: 162 ACAAATTTTGGGATTAAGCTGCTCCCAAGACAGTCTTCATCACCTTTGTGAAC  
TGGAAACACCAAATAGTCAAGGTCTGAATTTCCATTGTGTGTGGCTAAGACCAGTCGCATAGGTTN  
ATAAATGTAT

SEQ ID NO: 163 ACATATTGGCATTTCATCCTCAAAGGAATCATCAAAAAGAAAATTCAGTGAAT  
AATCTTTTTACCATGACTGTTGAAGTGAAGGGTCCCTATGAATACCTCACACTTGAAGACTATCCC  
TTGATGATTTTTTTCATGGTGATGTGTATTGTATATGTCCTGTTTGGTGTCTGTGGCTGGCATGGT  
CTGCCTGGCTACTGGAGAGATCTCCTGAGAATTTCAGTTTTGGATTGGTGCTGCATCTTCCCTGGGAA  
TGCTTGAGAAAGCTGTCTTCTATGCGGAATTTTCAAGATATCCGATACAAAGGAGAATCTGTCCAGG  
GTGCTTTGATCCTTGACAGAGCTGCTTTTCAAGCAGTGAACGCTCACTGGCTCGAACCCTGGTCATCA  
TAGTCAGTCTGGGATATGGCATCGTCAAGCCACGCCTTGGAGTCACTCTTCATAAGGTTGTAGTAG  
CAGGAGCCCTCTATCTTTTGTCTCTGGCATGGAAGGGGTCCTCANAGTTACTGGGTATTTTTCTTA  
TCCCTTGACTCTGATAGTAAACCTGCCCTCTCANCAATTGACGCCTGGGTTATTTTATGGATATTTA  
TTAGCCTGACTCAAACAATGAANCTNTTAAACTTNGGANGAACATTGNAAAACCTCTNTTGTTCG  
GCA

SEQ ID NO: 164 ACATATTGGCATTTCATCCTCAAAGGAATCATCAAAAAGAAAATTCAGTGAAT  
AATCTTTTTACCATGACTGTTGAAGTGAAGGGTCCCTATGAATACCTNACACTTGAAGACTATCCC  
TTGATGATTTTTTTNATGGNGATGNGNATTGNATATGTCCNGNTAGGNNTNAAANC GTTNGTTNG

NNCAACNNTTNTCTGGNGATGAACCCCTGNATTATCGCAC

SEQ ID NO: 165 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGANANGGAGTCTTGTTCTGTC  
ACCAGGCTGGAGTGCANAGGCACAATCTCGGCTCATTGCAACCTCTGCCTCCCGGGTTCAAGCGA  
TTCTCCTGCCTCAGCCTCCTAAGTAGCCAAANATTATAGGTGCCCCGCCACCACACCCAGTTAATTTT  
GGCACTTGTAGTAAANATGGGGTTTCACCATGTTGTCCAGGATGGTCTCGACCTCCTGACCTTGTG  
ATCCACTAGTCANAGTTTGTTTTTAAATGACTACCCGCGTACGCTTGGGAGTCTCAGCAGGGGGA  
TCATTACAGTGAGGACAGACACAGGTGAACCTATGGGTCTGGAACAAAAGTTATCCTACACCT  
GAAAGAANACCAAACCTGAGTACCTGCCCCGGCGNCGTTCNAAAGGGCGAAATTCANCCACTGGC  
GGCCGTTCTAATGGATCCCAACCTCGGACCAAACCTTGNGNAAAACATGGCATAACTGGTTNCTGG  
GGNGNAAA

SEQ ID NO: 166 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGTATATCCTGGCAATGGGATTGCTGG  
GTCAAATGGTAACTCTGTTTTAAGTTCCTTCGAGAAATCTCCAAACTGCTTTCCACAGTGGTGGAAC  
TAATTTACATTCCCACCAACTGTGTATAAGCATTCCCTTTTCTCAACATCCTCTACAGCTTCTGGTT  
TTTGTGTTGTTGTTGTTGTTTTTGAATCTTAGTAATAGCCATTCTGACTGTTGTGAGATGGTATCTC  
GTTGTGGTTTTGATTTGCATTTCTCTCATGATTACAGATGATGAGCCTGGCCATTTTTGATCATCTC  
ATTCCTTGCTCATTTTTAAATCCTTCTTTTAAAACTATTTAATAACAAAGTAGAATCAGTATCTGGG  
AATCCAACACCTGATGTCCCTTGGGGTGCCAAGAANTGTCTCAATCTACTGGTTGGTCTTTCTGN  
TTACTCTTGCTTATACTGACTTGGTTCTTNGANNGGGTGATAATTGAATGGNGAACCTCATGTGAA  
TGAAATGACTGCCGGAATCTGAGGAGCTGATTCTNGNTGGGTCTTAGAAGGATTGGTTGGTCTGAT  
TTGATNCNNGNTGTACNNCCTGGANCCTAAATACTTTAACTNGGANTNCCCTGATGCCCGNGGTA  
AANATCTANACCTGGCCGACCCTANGGN

SEQ ID NO: 167 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGAGGAGTCTTGCTCTGTTGC  
CCGGGCCGGAGTGCACTAGCATGATCTCGGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAATT  
CCCTGCCTCAGCCTTCCGCATACCTGGGATTACAGGTGCCCGCCACTACGCCCAGCTATTTTTTTG  
TATTTTAGTANAGACGGGTTTTATCATGTTGGCCAGGCTGGTCTCAAACCTACTGACCTTGTGATTC  
ACCCACCTGGGCCTCCCAATGTGGTGGGATTAGAGGCGTGAGCCACCGTGCCCGGCCAGTGATTC  
TTGTTAGAAGTGAACTTCAGAACATCCATCCACATGAGTGGAACATCATGAAGCAAGATGCTGG  
TTCTATCAAAGGAATCTTACATAGCGCAGCATTCAACATGTNATGAGAATAANAACCTCANACTC  
CCCCATCTNAAAAGTTGGAAATTATTAAGCCCTGTATGGGTGAAATNTGTTTCTAAAAATTAC  
CCGGAAGNTTAATNCCNGGAACCCCTTTTGGANGGGGGNAAAAACCCCTTTTCCAANTTAATTT  
TTTNAAGAGGCTTTCCAAAAAAGG

SEQ ID NO: 168 GGTACATCCCTGTTTATCCCATCCATCCACCGAGGCCCAACAGCATGGATGA  
TCTGTTTGACAGGGAAGCCTCCCTGCTCCCGTGACAGCTATCTACCAGCTGACACTTTACCATATC  
TGGCAACAAACTGTTTGCTCTCTTCTGGATTTCAAATCCACCAGCTTTTACCAGGGCCAGGGCCA  
GGCCTCCCCCATGCAGAAGATCTTCATTGGCTGCATTACCACAGCATCAACAGCATGTGTGGTGA  
GGTCATCTTTCCAGACTGATAACTCTATCCTAGGAGTCAGCATTCTTCTGAACACTTGCAGAGATTT  
GCTGNTGCCTTCTGAACTGGANAGACCAGGGTAGAGATCAGCCAACTTATTCTGGAGGACTTN  
ACACAGCTGACCTCATTATTTTTTAAATTTTGAAGTCNTTGNNGGTAATGGGAATTTGCCCACTA  
TAGGTTTCTTCAAGAACANCCATCTTTGANTTTTNTTGNAACTGNTGTTCCGNCNCCATGGAAAAN  
TCATCTCNNGNCCCGGGGGAGCTTNAANGGNTATNCCCTTTNGCTTNAANNTNANNGNNTTC  
AAATNNGGGCCCTTTNNGNGCGNTTTNNGGGAAAANNAANGGCCNNNNCTNNCNGGNGGNNTN  
AANGGGATNNTNANTGGGNGTTN

SEQ ID NO: 169 GGTACACACTCACATCTGGACCTGTGAGAACAAAAGGAGTCTGCCAGGATCT  
AAAATAAAGGCCAGGGAGAAGGTGCAGTTTCAGATACAGTGCATGGGCGCCACTGTGGGCTGG  
GTCAATGAGTGTATTTGGCAGTAACATGTATGTAAGAACTTAATCCACAGCTTGATATAAGGCAA  
AGGCTGATAAAGTCAGAACCAGCAATCAGAAAAATCATAAAGACCTGACTAGCCTGGGCAATATA  
GCGAGACCCCGTCTCTACAAAAACAGAAACAAAATCCCCATACATTAGCTGGGCATGGTGACG  
CATGTCTGTGGTCCAGCTACTTGGGAGGCTGAGGTGGGAGGATCGCTTGTGCTTGGGAGGTGA  
GGCTGCAATAAGTGGTGATCACATCGCGNACTTCCACCTGGGTGACCTTGTGAGACCCGCTTA  
AAAAATGTCTGGCGNCGGGCTCACACCTTTAATNCCACACTTTGGGANGGTTGAGGGGGGNGG  
TTCATNTGAAGTCAGGAGTTTGAAACCATCTTGGNCCACATTGGGGNAAACCATNTTTTTTTT  
NAAANNNNNNNNNNTNNNNNNNNGNTNCCCTTCCCGGGGGGCCCTT

SEQ ID NO: 170 CGAGGTACACTTTTCAACCAATCAAAAAACAACCTCTAAAAGATTCTATTA  
TGTAATTCAGTTTACATAAGTATATTTTTTAAATTTTGTCTCTCAAAGCACTTTTGGACTCTCA  
AATGTTTATTAATATGTTTACACAGATGAGGTTATTCCTTTAATGTTGTATCCTATGTGAAAA  
ATGGGGGTATTTGTACACAAGTAATTCGGTTTCTCCACATCCTTGCCAACATTTGGTTTCATGACTA

TTTTTIACTTTAACCATCCTGATAGGCGTGTGCTGGTATCTCACTGTGTTTTTAAAGTTTGCATTTTC  
CTGACCACTGATGATGTTGAACATCTTTTCACGTGCTTATTTGCCATCTGTATAACCTCTCTAGTGA  
AATGTCTGCCCCGCGTACCTGCCN

SEQ ID NO: 171  
AC1TTTTTTTTTTTTTTTTNTTTTTTTTGGGGGAAGTCCTATTTATCATTTTAAAG  
AACAGATTGNGCCTTTGGTGTCAATNTAAGAAATACTGTCTAACCCAAAGCCAAAATAGGTTTA  
AGTTTTGTGCCTTAAATTTAGGTCTAGGATCCATTTTAGACAATTACATAACNGNGCAAAACTGAAA  
TGTGTTTTTAAAAATATGAATAACCAATTATTTAAGCACTATTTGTTGAAAAAAATATCATTTTGA  
AAATCAGTTGTCCAAATNTATGTAGGTCTGTGTCTAGATTTGGTTACACTAATNTATCTTTATGCCA  
TACTATATTCCTCTTGATTATTGCAGGGTTTTTTAAAAAGCCTTGAAATCAGGAAAAGTTATGATTC  
CTAACTTTGTCCCTTACCAAGGGTTAGACTATCTNGNCCCTTAGTATTTAATTATATNANCTTGTTA  
NTTCTACAAAAAAGGTATNTTGGAAATTTGATGGACATTTGGCTGAATCTNTTTTCAACCTTTNCC  
NAACTGGTATTNAAATCTGGAATTCCTTTTCCCTGCAATCCGGTTNNTTTGGCCGACACCTTNG  
GGNAATTNACCCCTGNNNGCGNTANTNATNGGATNCANCNNGGACCNACTTGNGNNTTATGNNA  
NCTNTCCNGNGGNAANNTCCCNCCNG

SEQ ID NO: 172 CGAGGTACTTTTTTTTTTTTTTTTTTTTTTTGGGGTTTGATTCCCTTCATTCTC  
ATCATCAGAGGAAGGACCTGCCAACGTTGACACTTGCTAATTACCAAGTCTGGCTAGCTGATC  
CACCCGCGT

SEQ ID NO: 173 GGIAC(TTTTTTTTTTTTTTTTTTTTTTTTGGGTAGAGACAATGTTTCACTATGT  
TTGCCAGGCTGGTCTCGAACTCCTGACCTCAGGCGATCCACCTGCCTCAGCCTCCCAAAGTGGTGG  
GATTACAGGCATGAGCCACCATGCCTGGCCCACTACTGAGATCTTATCCGGAAGTTGCTGATTAC  
CAGCTTCAGGTGTTTCTGTTTATTGGGAGACTGTTCTGCTGCTGGCTGTGACCAATTATTATTTTA  
AAAAGACAGTTAACTGCCGGACCATCATCTGATGGTTGCCCTGACATCTGACATTCCTGTTGTG  
TGTGTTTTGAGGNGAGGGAGCCCTCTCCTGCCCTGTTCTTGTCTGACTAGCTACCTACTGTAACAA  
AACTATAATTTGGATTCCATAACGTGATACTCAAATGAAATTTCAAAATCTTTCAAGACATTTATGA  
ATCATCAACTTTCGGGGTTGGTCTGTAGGCTCACTTTAATGCAAGTTCTANCCAGGGCAGCANAA  
AACTATTTGAATTCGGNGNGCCCTNTTAAANAAGGATTTNTTNAACGTTCAAACAGGGCAGTCTGGC  
CNTTGAAAAAAAATTTCNTAGGTAGGGGTTTCATCNTCCAAATACTGCACATTTGGTCAAAAAAC  
ATTTCAATTTNTTCTAACCCCTGNTAT

SEQ ID NO: 174      ACCCGGGGGGCANNCCNGTGGTCCCATAGCACAAAGCTGTGAGGGGATTCACTT  
GTGTGCNGAACTCCTCGGAACNTGGTGTCCCTAAACATNTTCCTGGGAACAGCCNTNNCTAAGA  
CCCTGATGACTANNAGCTANCTAAGATCAGCTGANTTA

SEQ ID NO: 175 ACTTTTTTTTTTTTTTTTTTTTTTNTGAGTGAGGCAGGAGTCCAANGAGGNTAT  
TTGNGGCANTAAAATTGATTAAGGATNCTNGTTTANGANATCAGGTACGTCCTTLAGNGTNGCGT  
ATGGNTATCANTCGAATTGAGGTTA

SEQ ID NO: 176      ACTGGGATTACAGGCATGAGCCACTGCGCCTGGCCCANAAATCTCTTTTGAA  
CANTNTTCAAAAAATACAGCTAGCCTCAGTGGTTCATGCCTGTAATCCTAGCACTTTGGGAGACCA  
AGGCAGGCTGATGGTTTGAGGGCAGGAGTTTGAGACCAGCCTGGGCAACATGGCAAAACCCCATT  
TCTATTANAAAAAANAAAAAACCCTTGGCATGGTTGCACGTGTCTGTAGTTCANCAA  
CTTGGGAGGCTGAGGTGAGAGGATCACCTGANCCCAGGAGGTGGAGGCTGAACTGANCTGTGATT  
CCGCACTGCGCTCCAGCCTGGGCAATANACAAGACCCTGCCTCACCAACCCCCCAAATACCATT  
TATAATAACTTTAAAAAAGANTCGCTTAGGTGAAATCTAAGGACTCATATAGTGAAAGCTATA  
AAACCCTGATGGAAAACATTAAAGGAGACCTAATTAAGGAGAGACACACTGCATTTGNGGATT  
GGAAGACTGGACGGAGTAAAGACCTCNTTCTCTCCATATTGATATGTAAACTTANTGCAATTCCTA  
TTCAANCTCAGCAAGACTTTTCAGATATAGACAAGATTNTTCTACANTTNACTGAACGCTAAGG  
GAACTANAATAA

SEQ ID NO: 177 TATACCACTCACTATGGGCGAATTCGAGCTCGTACCCGGGGATCCTCTAAGTC  
ACCTGCAGCATGCAAGCTTGAGTATTCTATATGTACCTAAATNCNCCGNGAAGAAGGCNGTTTT  
GCGTTATTGGGCGCTTCTTTCCNNCTTACCTCGATCAACTTGACTCENNNTTGCGCTCGGGTCNTTCN  
NGCCTGCGGCNGAAGCCGGTATACANGNTCACTCAAAAANGCG

SEQ ID NO: 178      ACTTTTTTTTTTTTTTTTTTTTTTTTGC TTTTTTTTTTTTTTTTTTTTTTTTTTTTNG  
CTGGNTAAAAAAATTTGGGT TTNATTGNTTTNGCTAAATAATACTAAAAAAAATTTNATTTTNA  
AGGCAGGGCTTGAATTTNTTAATNGATCCATTTNTTAATTA AAAAATAAAGGAGNGGGGAAAGA  
NATCATGGCCAAAAAAATTTGNNTTAACCCCNCCCCCCCCAAAANGTTTNGCNGTNTTGGAA  
CCATNNCCCN TCCNCNCAAGGGGCGCTNANTTTCGGATGGGGGNAACCTTNC CCAANA AAAC TGCC

TNAAGGNCGGGGCAATGGGTGCCAATTTTACCTNTCAGCAGGTTAGTCAACCANACAAACNNGGGG  
GGCTAAAGTCCAAAAATTTCTTCCAGGTTTTCNTGCTTATTTGGCTTGANCACATNCAAACGTGNCN  
TAANCCTGTAAAATTTAAGGGGAGTTGGGGTGGGGGCGTAANAGCAAANGGACAGCCGGANAAN  
AGAAATNNCGGGTCCCCCAAGTTTTTCTCGGNTAGNGGCTTTNGNTTTAAATTTAAANNANA  
GGCCAGAGTAAATGGCACNCNCNNGTTTTTTTATNAAAGAAAAATTTTCCNTAATGAGGGAAC  
NTNAA

SEQ ID NO: 179 ACAGGCTTGAACAGAAATTGGAGAATGCCTTGAAAGACAATAGAAAAGTGCC  
CACCCACCAGACAGACCAACTTGAANGGAGCTTTATTGGCCAAGTGGTATNCCCCTTGGGAACCT  
TTTATGAATGCTTANNCAGTGGANTANAGATCTNCTTCCGAAACNTCCCAAANATGANTTTTGATN  
AGGGGAAGGAAAAACCAACCTTTCANNATTGNTTGGAACTNAAAACAATTTNGGAAAAAANGGG  
TTTCCANANAAACCTTGGGCCTCCANAAAGGGCACATTNNACCTTGGGTTCAAAACCTTTTTTTTC  
CCTTNTTTGGGCCAAGGCCCTTTAACCCNGCCCT

SEQ ID NO: 180 CCCCCGCGGGCACCTGGAGCAGAGGGGTAATGACCACTGGAAACACTTGCG  
GTATTCCAGGAGAACTTCCCCACACATCGCGGGCAACAGGGCACATGCTGAAAACATAATGCCCAA  
AATCTCTCGAAGAGCACCTCCAAACGCTAATCAGATTCCCAAGTGAATGAGGAGACATGTGATGG  
CACCGAGAGAGAACTGGTCGGCTTACCCTCAAGGAAGCCATCAGACTAACAGCCGATCTCTCTGG  
AAGAACCCTACAACAGAAAGAGTGGGGCCAATATTCAACTTCTTAAAAAAGAATTTCAACCCAG  
AATTTATATCAGCCGACTAAGCCTTATAGGAAGGAGATATGTCTTTTCAACAGCAAATGCTGAGA  
GATTTTGTCCCCAGGCCTGCCCTAAAAAGAGCTCTGAAAGGAACACTCTTGGAGGACTCCGTGT  
CCTTGCCNTCACACTGGNCGGGCCGTTNCTTAGNGGNTCCNGANCTCNGTACCANGCTTGGCNGA  
AATGATT

SEQ ID NO: 181 ACTNTTTTTTTTTTTTTTTTTTTTTTTTTTCTAAAAACCACCTTCTGAATTTGTTGAT  
CTTTTGAAGGGTTTTTCGNGTCTCTATCTTCTTCAGTTTCAGCTCTGATTTTGGTTATTTGTTGNCCTCT  
GCTANCTTAGGGATTTGTTTGCTCCTGGTTCTCAAGTTCCTTTAGTTGTGATTTTAGGTTGTTAAGT  
TGANATCTTTCTAACTTTTTTCATGTGGCCATTTAGTGCTATAAACTTCCCTCCTAAAACTGTGTTGG  
CTGTGTTCCAAAGATTCTGGNATGTTGTATCTTTTTTCTGATTAATTTCAAANAACCTTCTGATTTCT  
TGCCTTAATTTTCATTATTTACCCNAAAGTCAATTCAGANCAGGTTATTCAATTTTCATGCTTTATTA  
GCATGTGGAATATTTNGGAACAACTCCAAAAATGCTCCTTTTGTGTGGGGTGTTTTATTTTTTGCAT  
TTTTATTATGAGAAAATATACATGACATAAAATATANCATTTAAAAATTACTAAATATATATTTAA  
ATGGCATTTAAGTAAATTCACAATGTTGNGCAACCATCACCATTATATATTTCCANACTTTTACA  
TCATTCCAAACANAACTTTTGNCT

SEQ ID NO: 182 ACGCGGGGAGAAGTTAGGGGCTGCAGCGGCGCTGGCTTTAGGTGAACGACGT  
GAAAATTACTTTTCCCACTGAAACACACCCCAAGTATATGCCAGCCTTCATGAAAGTGAACAGAG  
AAACGAAGCGCCTTTATGTGGGTGGCCTTAGCCAGGACATTTCTGAGGCAGACCTACAAAATCAG  
TTCAGCAGATTTGGAGAAGTTTCGGATGTGGAGATCATCACACGGAAAGATGACCAAGGAAACCC  
ACAGAAAGTTTTTGCATATATCAACATCAGTGTAGCAGAAGCGGACCTGAAAAAATGTATGTCTG  
TTTTAAATAAAACAAAATGGAAGGTGGAACATTACAAATTCAACTAGCAAAAGAAAGCTTTCTG  
CACAGATTGGCCCAAGAGAGAGAAGCTGCAAAAGCTAAGAAAGAAGAATCAACAACAGGTAACG  
CCAACCTGTAGAAAAGACAGGAGGAGTGGAATTTCCATATGAAAGCTGTGCCAGGGACAGAAGTT  
CCATGGCATAAGATTTGGGTTTGTGAGCNAATTTGGGAAGAGTCTTACCTGTTTCTTCACCTTAAA  
AATCAACATANNCGTAAAAATCATCTNATATGGATCCCTCAAAATTCTGCCCAACCTG

SEQ ID NO: 183 ACGCGGGTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAACT  
TAGCTGGGCGTGGTGGTGTATGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAAGAATCAG  
CTGAATCCATGAAGTGGAGGTTGCAGTGAGCTAAGATCACACCACTGCACTCCAGCCTGGGTGAC  
AGAGCAAGATTCCACCTCAAAAACAAAAACAAAAAACAACCCAAAAAATAAAATAAGTAA  
ATAAATAAATAAAGGTGGAGTGACTATATTAACACCAAAGGTCTTTCCAGGACANGTATCACCAG  
ANATAAAGAGGGTNATTTTCATANAGGCAAAGAGGTCAAGTGATCCAGAAGACACCAATCCTAAG  
TGTATAAGTAACTAATANCAGATCTTCAAATACGTGATNTAAANGCTAATAGAAGTGAAGGAT  
AATTAGACAAATCCAAATTAAGTTAGATATTTCAATACCCCTTCTNAAATTACTTATAGAATAA  
GTAATCTGAATATTTANNATTA

SEQ ID NO: 184 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTNATTTTTTTTTTTTTTTTTTTTTTTT  
TTTTTTTGGCAGTTTCTAAGNCATTACTTTTNATTTTGAAGGATTTGGGAACTNTTNANATNANG  
GGGANAGTTTGTATGATTAATAAAAAANCAGCTTTTNNATNAAATGCTNNGNGGNGAACGAGTTNT  
CAGCCTGNGANATCCGACCATCCCATTAACCTTTGAAGNTTNTCTNGATTAATAAAAAAAGG  
GGNGGNGGAAAAAAGGNGGAACATGCTAAAAACCTAAATGACAATCATCCAAATGNGAGGAAA  
NAANAACCGATTNACCAACTNCCTTTTNTNTTTNAAAACCTTCTANATCTNANTNTATGATTNG

GCCTTCCTGGCTNAAAAAGCCTGCAGNCCCANAGAACCCNTGAAAANAGCCATGGNTCTNCAAAA  
GAANTAGGA

SEQ ID NO: 185 ACAGTATTNTGAATGTGAGATGATTTGTCAGGACTAACTGTCTTTTAAACAAA  
ACATTTTCAGTNTTTTAAATAAAATTTTGNAAAGNAATGTGAATTAATAAATCCTGGAACANATNTG  
AATTCATTCACTATTGNGTANGAANATGCTGTAAANACATAGGAAGGG

SEQ ID NO: 186 ACTTTTTTTTTTTTTCTTTTTTGTTTTTTGGTGATGTGGCTTAAATGCAATAGTT  
TCTTTTTTGGGACATATTTCTGCCAATTAAGACTAGAAGGGCACAACCTTTTTTTTAAATTACCATA  
GAGAAANATACATTAAAAAAAATCTTCTGATGTTTTGTAGCCATAACTAAATTATGGTAAAAATGTG  
CACTATTGTGAAAAGGAGCAACGNAGTTTTGGGTTTTTGTGTTNGTTTGCTTTGCTNTGTTTCAT  
AAGAGATTAATAANGTTTCTGGATAAGGGATTAGCTTCTCGAAGTGTCCATCATTCTGNGTAANAA  
NCTTAAATATGNAATGTACCAAACCTCCANAATTAATAAANCTCTCATGTTGTTATNCTTTATACAA  
AGCAATGATAACNGCNTATAACACTGGCATTNCATGGCAAATGCTGCTACCTTNGTTTANANAA  
CAATNTTCAANCAAAAANGACTTGCTNTAANGTGNTTTTAAANANGCAANTGGTCCNAGANTTTCT  
TTANAATGANAGTTTTATTGNACTNCCCTTTTTCNANNTGGTCTNATTTNGCATTTGNGNTNCATNC  
CCTGNNNGGCGTTNNGANNNGGCTAATTCATNNNACNTGACNGCCGTATATGGTGGACCG

SEQ ID NO: 187 ACTTTTTTTTTNTTTTTTTTTTTTTTTNGTAATAAACATGTGGAGNGTCNTGTAN  
TCAAANAAAGCGCANCAAGAGAGGCATTGCTTGCTGGAACACTTGACTCCAGTCATGTGTCAAATC  
ACCTNTACATNTAGCTTTACCAANAATGNATGCTAATGTATCCTCTTTTTTATTTAATTNANCTGANT  
NGAGNNNTGGCTAATATTAACCAACATGNAACCAACTAAATGTAAACACCATCNCCAATTTGC  
AAAACATCACTACCAAAATTCNANTATCCTAANTTTCANGTNCCGGCCCGGCGCCTTTCNAAAG  
GGCNAATTTAGCACNTTTGCTGCCNATCANTAGGGCCTAGNTA

SEQ ID NO: 188 ACCACTTCACTCCAGCCTGGCGACAGAGTGGAACCTCCGTCTCAAAAAATAAA  
ATAAANTAAANTAAAGCNAAAATNTAAANTGTTAAAAAAAACAAAAAAGGGAAAAAGGANGC  
TGATTGCCCTTGGTGAGTCAACACTGGGTATTTCTGACCACTATTTGAAACAAAAAAGGAAACNAC  
TGATATTCTATGCAAAGATCTNTTCTGGANGGCACNTTGCNGNACACCAGTGNGNACTNTGAT  
NANCCCTTCATTGATTTGAAT

SEQ ID NO: 189 ACCTCCAAAGTGTTAAATAAATTAAATTACCACTGGAAGAGAATAAAAATT  
TTAGTTGATCCACATTCTCAATGACACCTGAATTTCTGTTTTGTTTTGTTTTCTAATCTGGTGGGAC  
ATAGCATGGTATTTTCATTGTCTGAATTTATATTTATCTGAATACCAGAGTAGATGGGGATCTCTCAT  
ACTTTTTTACATTTGGCTTCTTTTTCTGAGAATAACTTTTTTATTATCTTTGTCAGTTTTGGTTTAGG  
ATTTTTTTTTAGTTAAAGGGCTTTTTAATAGTTACCATCAGAGTGGGAACAAAAAATAAAGTTT  
ACAGTTCGGGGCTGTTATTTTAAATTCCTTTCTTTGCACTTTTCTGTTCTATGCAAAGTGTTGCTGTT  
TGCTCACCCGGTNAAAGGAAAGTGANTTAANCAATTACTCTNTACTTTCACAGCCTAAGAAAAAT  
GTNGGAGGTNAGAACNAAGCTTTNATTCTG

SEQ ID NO: 190 ACCCTACCACTGTTGGACCAGTGAGAGCAGTGGAATTGAGATCTCGCTACCG  
TTCTTCACCTACCGTCTACAACCTACCTACTGACAAAGAAGACTACATGACCGACCTACGAACCTT  
GGATACTTTTCTCAGAAGTGAAGAGGAGAAACAGCATAGGGTTAAGCTGGGGAGCCAGATTCTA  
CCTCTCCTTCCAGCAGTCCTACTTTCTGGAACATAAGTCGTTCTATGGGGGATTATGCACAACTTT  
AAAGAAGTTTCAGTATCAGCTTGCTGTAGGTCTCAGGCCCATGTGCTAACAAAGATGAAGCCG  
ATCTCAGCTCTAAACAAGCCGCAGAAGAGGTCTGGGCAAGAGTGGCTATGAATAGACAACCTTCTT  
GATCATATGGATTTCATGGACAGCTAATTTAGAAATTGGATCAATGAGACAATATAGTGCTTGTTC  
AAANATGAATCTGCACACCANATGAAACAATGGNTGGCCAACCTAAGATGGAAAGGTNGTTTCTA  
CTGAAAAGCTGCCTGGTAAGNNCTTTTTTCANTT

SEQ ID NO: 191 ACACAATCCTTTATAAAAGTNGTATATATTTTTTTTCTGTCAATACCTTCATTAC  
ATTTATAAATACAAGATTTACACAGCACCCCATCAAAAAAATTAACCCTTTACAAATATCTA  
CATATATTTACATCCCTATAAACTTTCAAAGGGGGTGCTCTGGTNAAGGNCNCCCTAGTTAATGGG  
CCATTTACTGGGNGCAGCAAAATACATATTAATCNGNAAAGTTTTTTTTTGGCCATAAAACATTTT  
NANAAAATNAGTCNAGANACTTTTTCGGGGNAAACNCANTNAGNCTNATAAAAAACCTTNTTTTA  
ANAAAGGGAACCTTAAATGGGGCGATTTCCCTGCCGAATCAAAAAGTTAAACCAGATTTGGGAAT  
AAGGTTGATTTGGGAACCAAATGGCTCTNCNAAACCATTTAANGGCAGGGANTGNGAATGGGTNG  
GAAAAAACCTTGGGAAAGCGGANATNGGAANANCTGGGGANGAAACCAATTNTTATTATGCCCTT  
TTGNGTTTGAACCTGGGNTTAAACTTCCCNCGTTTCAAGNTACTAANNNTTATTNCCTTAAGGAA  
TTTCCAAATTGNNAATNAACNTTTTTTAAACCANNCTCAATTTTCTTTNNGGCCCCCAAAATGCT  
TGGAANNNGNTTTTTTTTTT

SEQ ID NO: 192 ACTCTGGCTTGTGCTTAATACTGTGGTTAAGAGAATCACCTATATTGTTGCAT  
ACAGTTTATTCATTTTCATTGCTGTGAAGTATTATATCCAAATATTATAATTAATTGATCCATTCTA  
TTACAGATGATCATTTTGGGTATTCTAGTTTGGGGCTATAATTTGTGCTGCTGTGAATATTTCTCAA  
ACATGTTTTTTAAAAATATATTATGCATTTCTGTTTGGCATTATGCCTAGAAGTGTAATTTCTGGG  
TCATAGAGTAGCATTATTTCTAATTTCTAATGCTCTCANAGGGAATACNNNGGAAATTTTAGNAAAG  
CAAAGGNCNTNTTTAAAAATCTGTAAAGGAAAACTGAATTTGACATTGTATTTCTGNCCCTTNGACA  
ACATGGCAATAAACATTTAAGCACTTTTCNATATTAAGACAGAAAAAGCAATGAATAGCAGANGN  
GAAAATGGAATTTCCCCCAATATTTAAGTAAAAAAGAAGTNTGATGAAAGAAAAGTTGGNTTTTTA  
AGNGGTTTTATTNTAGGAATTTTAAAAATATAGGATATTGGAACATTCAGGTANGATGCTNCTTA  
ATTTTTTTTNTCCNGNAATTTTNGGGACCCCGANNTTAAAAAATTNGAAAGACCAACTTCAGGGN  
GGNCAATTAGAAAAAGG

SEQ ID NO: 193 GGTACAAAATAAACTTTGAGGCAAAAGGCATTGCTGCAGATAAAAAACATGC  
CTATATAATGACAAAAGGCTAAATTTACCAGAAAGCTATATCAATTCCAACGTGATTATACCTGAT  
AACACAGCTTTAAGAAATATGAAGAAACATGACAGAATTAAGGAGAAATAAACAAATCCACTGT  
CATAGTTTGAAAATTTAACTAACCTTTCTCAGAACCTTATAAAGCAAGAAACCACTAGATTATTTT  
CAATATCATCTCGATTTATAATCACAATAAGTTTCCCTTTTGGGAATTAATAGCAATTTATTTAGT  
TTATATGGATAAAATAAATACCCACAAATTGGTTTTGAAAACTATTTAAAAAAGGAGAACTGAG  
AAGAAATACTTGCTGTAGCAGATACTGAAACAACATTCTCTAAGGCCATTTTACTGAACTCATTCT  
AATAACTTTATGCGATCAGTGTAATATAGTAGAAGTTCCAAAATTAATATGAAAAATACATGTC  
AAAAATATGAGAGTTCAATTTAAAAGGAAATAGTGACTAGTTAATAAATGATG

SEQ ID NO: 194 GGTACTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGGGTCTCACTCTATCACCCAG  
GCTAAAGTGCAATGGCGTGATCACGGCTTACTACAGCCTCGGCCTCCTGGGCTCAAGCAATCCTCC  
TGCTCAGCCTCCCATGTAGCTGGGACCACAAGCATGCNCCACCATGCTCAGCTAATTTTTTAACT  
TTTTGTAGCAACAGGGTCTCACTTTGTTGCCAGGCTGATCTCGAATTCCTAGGCTCAAGCATTNT  
CCCACCTTGACCTCCCANAGNGCTGGGATTATAGGCATGAGCCATTGCACTGAGCTCACTAGCCA  
NAATTTCTAAAAATCTTCTNTCAGGAGACTATANATAATGNCCTACTTCTTCTCCTTTCCAGTATCA  
CAAATAAATGGGCTGNGGGTAANCACATTTCTTTTCTTTAATCTATTTTGGNGGGGCATTGCATT  
TTCTATGTGGGAACCTAAANANNTTATTTCTTAAATATAAAAAAGGTTATAAAATTTACCAAGCTNT  
TGGGGGGGNGGGAATTTTTNTTTAAATCACCGGCTTTNG

SEQ ID NO: 195 ACTTGATNNGATTCTCAGCTTGTTGCTGTTGGTGTATAGCANAACCTACCGAG  
TTGTGTAGGTNAATCTAGTATCCTGATACTNTGCTGAATNCATTTACCAGTTCTAGGAGCTNTTTA  
GAGGANTCGTTCGGGTTTTCTANATNTACNCCGAAGGAGGGAGGNAGGACAT

SEQ ID NO: 196 ACTGCGGGGTCCTTGATGGACCCTAAAAGGGGTTGGAGAGACCGATTACAG  
AGAATGATTCCATCAGGGGAAATNCTCCATGACTGGCTTTGAANATGGTGGGGGCCATGTGACAA  
GGAGTGANGATGGCCTNTAGNAGCTGANAGTGGCCCCCAGNCNACAGCCTGNNTGAACTAGNAA  
NCNCATTATTTCTCTGAGCCTCCAGAAAGGAGGAAAGCCTAGCCAACNCCTTAATTTTAGCCTTGC  
AANATGCTGAGCAGAAAACCCAGAAACTNCTGGTTAACGTGGNGAAACCCNTCTTTACTAAAAA  
ATACAAAANATTTTGCTNGNCTCNGGTGGNTNGTGCGCCTTGTANTTCCCATCTACTTTNGNAAG  
CTTGACGGNAATGATAAATGGCNTTAACCC

SEQ ID NO: 197 ACTTTTTTTTTTTTTTTTTTTTTTTTNGGCCAAGCATTTTTTAATAAAATAAAAAAC  
AAATGAAACCCCANATTTAATTAATAAAATTTCCCATATTCTGGCCTACTCTGTAATTTTTATGTTGC  
TGCTGAAAGGNATTATGTAGTTACTTAATAAGANAGAGGAAGGGGAAGTATTTTTAATTTTAA  
GGATGAGAGGAAAGGAAAACCCTTTGGGTTTTNTTTTTNTGCGTCTGGAGCCCAACAAGCGGC  
AGCTGCGTATTTCTTAGCAGTTTTCTTAGAGACAACTAACACAGCTTATCAATTTCTTTACA  
TCCTAACTCTACAATGT

SEQ ID NO: 198 GGCTGCANAACAAATCAAGCACATCCTTGCTAATTTCAAAAACTACCAGTTC  
TTTATTGGTGAAAACATGAATNCAGATGGCATGGGTGNTNTATTGGACTACCTGTGANGATGNGT  
NTGANCCNATTTAATNATTATT

SEQ ID NO: 199 CTAATTTATATGTTGCTCTGCTTATTAAATAATCAGCTTAAGGATAATGGGGT  
ATTATTCTACTCTTGTGATCCACGCGGGTAATGTNTNTNTTTTNCNAAACANATGNTAAAAATAAGA  
CCAAGATGANTTAATTTNTANNAGTCCTTATTACACAATAAACANTNNCTTAANNNTGNCAACTGT  
GGCCGAAACACGNGGCTCATGCCTATAANCCGACACTTTGGGAGGCCGTAAANGGTTGGATCAN  
GANGTNAGGAGATCCNAGACCATCCNGGCCAACATGGGTGAAACCCCTGTGCTCA

SEQ ID NO: 200 ACTTTTTTTTTTTTTTTTTTTTTTTTACAAGGGTAGCAAAAAATATNTGTAA  
NGCAACATTAATCAGTAAAAACTAGGAGNGGCTNTNTCAATACAAGACAAGATANACGTCAGT  
GGAAAAAAATTACAAGGCNCATTACATAATGACAAAAGGACTAATCTGTAAAAANATAAAGC  
AATCTGAATTCCAAACAGCAACAGCAGAAGCTTCAAAATACATGAAATGAAAACCAGTAAAGCT  
GAAGTAANAATAGACNCNTCTTTATTCATAGTTGAGCATGTTTACAATCCACTGNCAGGCAATTGA  
TANAAGTNTTGGCATATAAACAAAGAAAGGTTTNNAAAGAACTGGNCAACTTTATTGGTCCAGAN  
GNNTGTNNATAAAATTTATGGGATACATCNNCTNCAAANAAAAANTTTNCNCCTGGTCAAGANCTC  
AGAAANATTTTNTAGATGAGAACANGATGACCNTGNGAATTTTAAANAATTGCATTATGNAGAG  
TTTGATNTTTGCT

SEQ ID NO: 201 CGTACCAGATCCCACCTAGGGGCGCNACTTGCTTGCATAACTCCTAAAAANAC  
CTGGNCACCCNTNNNANCCNTAGGACNCTGACTNNNNGCTGGNAAAGGGGNTTNAITNANTTGNGT  
NNNAGAGTCAAACATATAAATTACNTNCCAAGGTTAGGTTCTACCTATGCCAGNAATGAACAAGG  
ACAGCTTAATAGGTTATAANCAAGATGGAGTCNTTNGGGTCTGATCTCTTTCAGTGCATAATTT  
CCTCAGTTACAATTTTGTAAAGGTGGNTTCAAATGCTTTGCTGACCTCCCATTAACAAGGATGTG  
CCGATTGGAACCTTNTNTTTTGC

SEQ ID NO: 202 ACTTGCTAGGTATCCTGGGTCACTGGCGGTGCAAACCTGGTTTCCTCAGCTGCC  
TGCCATGGGGCTGAGTCGTCAGGGACTGGTAGTGGCTTTGGAGCTNATAGCGGAGATTGACTGAA  
AATAACTGCAATTTACTGNCAAGCCTTCCCCTGAATGTTACAAGCCTTCAATANACTCCAGAGTTC  
CNAATAATTACATCAGACAAATCTGCCAGNGTAATTTTTGTCTGGGTGGGGTAAAGAATTCCTC  
GTGCTTCTACTCCAGCAT

SEQ ID NO: 203 ACAGTCATTTTAATGATGTTGATTCTTCCAAACAATGATCATGGGATATTTTT  
CCACTTACTTCTGTATGTAGTATTTCTTTCAGCAGTGTTTTATAGTTCTTGTGTGGACATCTTTTA  
CCTCATTGGTCTCCTTTGTAAATATATTTCTAGATATTTTATTTTTTGTGTCTATTATAAATGAAAT  
TGAGATTTTTATTTGGTTCTCAGTTTGAGTGTGTCGATCTATATAAATGAAACTGATTTTCAAGCA  
TTTATTTTGTATCCTGAAATCTTACTGAAGTAGTTTATCAGGTAAAAGAATCTTTTGGAGGAAATA  
TCAAGGTTGTCTAGGTGTAAGATTGTCATTAGCAAACAGAGATAATTTGACTTCCCTCTTTTCAAT  
ATGGAAGACATTTATTTATTTTCTTGCCTGATTGCTGAGAGTGGCCATCCTTGCTTGTTACAGTTC  
CCATTAGTATGATATTAGTTGTGGTTTGTATAGACGGCTCTTATTATTTTGCAGTATGTTTTTTT  
CAATGCCTAGTTTGGTGAGGGTTTTTATCATACAGACATATTGGATATTATTTGAATGTTTTTCTGC  
AGGTATTGGAGATGATCATATGGCTCTGTTTTTAAATTTGNTTCTGNGGGGAATCACATTTATTGA  
TTTGC

SEQ ID NO: 204 ACGTGACAGAGCCAGGCTTAAACGCAGATCATCTGGCTTCAGACTTTCATCA  
CTTTATTTAAATAGCTCATAAGAATACTATGAGGCTCAAATGAGGCTGGCGGAAACCACAACATA  
TGATATTAGTTTCAAAGAGTCATAACAGAAATAACGAAAACCATGAGGATGAAAAGAAAAGC  
CTTGTCTTCTTCCACTGTTGAGTTTTTCAAAGCATTCTTAGAAATTCGACTTCTCAAGTCATGAA  
TGTGAAATGGAGGCCAGAAATTTGAAGAGTTGAAGGCTGGTGCAATCACTTTGGAAAATATGAA  
AGTATTATACACATTCTAATTATATTTTTTGTGTCAGTCGTAAGATTAGACAACAATTTGAAGATGA  
CACACCGCTCATATTAGATAGATGTCTGTAGGGGGAATACTCCTTCCCCTGACAAGACCACATCGC  
CAGTAACGTCCTCTACACACACAGTTGGCCTCTGTGTGTCTCCGTCTTAAGAGTAATTCAGGACT  
ACTAGCCAAGTGGTTGGGATTTAGGAATAGAGTGGAATTCAGCTTACCTTGTAAGAACTAGGACG  
AGATAAAGCCTTTTCTAGCATATAGGCATTGNTGGCATAATCCAGCTCACTACAGNCTNACCCCCC  
GGCTTANGTGATCCTCCACCT

SEQ ID NO: 205 ACTTTTTTTTTTTTTTTTTTTTTTTTGGGATGGAGGGCCGCTCTGTTGTCCAGG  
CTGAAGTGCACTGGCATGATCCCAGCCCATTTGCAACCTGTGCCTCCCGGGCTCAAGCCATTCTCCT  
GCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCTGACACCATGACTGGATAATTTTGTATTTT  
TAGTANAGATGGGGTTTACCATTGTTGGCCAGGCTGGGCTTGACTCTTANCTCAATGNAGGAAATT  
GGTTGNAATGCAANCTTTAAAACTTAAACCCAATAAANGAATT

SEQ ID NO: 206 ACCGNATGCTGGCNGGAGGTGGCATATAGCTCACTGGNACTGANGGGCTGG  
GCACCCAACCCNTTCCACCTGTGCTAATCGCCTGGATCTATCATNANTGCAAAAANCTNCTTTTG  
TTGACTGG

SEQ ID NO: 207 ACAGAGAAACCCTTAGGCCAAACTTAAATATGTAAGGAGGCAGCTTTAGGC  
TAAACCTGATTTAACAAGGTGAACAAACAGAAGACTCTGGAGAAGTATTTCAAAGGCTACTGAGA  
GAAGCAGGAGTAGATTTAATCCTGTATTTCTCCTTTTCAAGAGATGTAAGTAGTGCCAATAGGCAGG  
AATTAAGAAAGAAATTAGATTTTTCTCAATTAGAATTTCTTAAATATTAGAAGTATTTGAAAATGAAA  
TTAAGTTGAGTGAGTTCCCTGTCCCTGGAGATAACCAAACTAAGACTGATGCCTACTTTTAAAG



AAGTCTCTGAAGATCTGTGTATTTCAGTAGAGACTTGAATTAGGTACTTCTAAGGCTNCTCTTTGNG  
ATTCTATGAAATTACGATTCCGGGAGAACCCCTTCCAAAAAAGTCTCTNG  
GCCGGACNCNCTANGNCNATTCCACNNNTGNGGCGTNATATGGTCNACTCGGACAANTTGNNGAC  
ATGGANACTGTCTTGGGAATNGTNTCGTCAAATCCCCANATCCNACCGANCTAAGGNAANCCGGG  
GGCTAAGNAGNCTACTACATTATGGTTGGCTNTGC

SEQ ID NO: 208 ACTGNNTGCAACAACTCATGGANTTTGATGGGGAAGACCTGGTCTCAAATAC  
CAAGGGGGGTCTGGAGCTNCCTGTGGNTTAGGNGGNGAAAAAANATNGGGAGANAGCCNGGC  
CAAGTTTTAGAACCTCTGCAAGCTCATGAAAAAATCTTATATAAAAAGGTGAGAAAGTGACANT  
CTNCATAAACTTGGGTCTTTCNCCTTGCTGCANTTTGACCACCACCTACCGGTGGACAACCCAAAT  
NTTGGAGGCNGGTTTCATGAAANCCCCNNGCACTTTTGGGACAACCTCCACCATNGGCTNTTNTNA  
TGGGCCNAAAAANCCCCCTGNAGAATCAACCCCTGCCCCCCTTTTGGANACNCTTNNGGCAA  
AAAGGTTTATGGNCCACAAAAATNATNNNGCCATTTAANGACCTNGTGGGTGCTTCTTTTNTNA  
AACCNCCTGNNTTTNTTNTGGCTTTTTCCTTGNGGATCCCCAAACCCNCTTCCAACCCGTNTTTTT  
NCCANTGAAANAAACCTNGGTTTANGTTTTTGAAAAAAANAATAATGGGGCNCCAAANGAACCC  
CCATGCTTCATTTCTTNGAAGAAAACCCCNCTTNGANGGCAAANAAGNATNCCTNTCCATTGGA  
AAAAANCCCTTTTGTTTANG

SEQ ID NO: 209 ACGCGGGGAAACGGAAGTGAGCGGGGGTCTGACTGACGGTAACGGGGCA  
GAGAGGCTGTTTCGAGAGCTGCGGAAGATGAATGCCANAGGACTTGGATCTGAGCTAAAGGACA  
GTATTCCATTACTGACNTTTAACCAAGGTGGACCTTTGGAAAGTCATGATCNTCTCGGAAAGGTTT  
TCTTGGGTGAAAAATGACCTTTGCCTAGCCTTCCCTTGGATTATAAAAAAATTTTNCNCCTAAC  
CAAGANAAAATGAATTTTCCCACTGNNAACATTCANGGTCTATTTGCTCCCTAAAATTCCGATG  
GAATTCAAGGCNTGCNCAGGTTACGGCTTCATTTTTTAAAGCTCAAATTTTCTGGGATTTTGAG  
GGNAATGATAAACTTTTGGTTTAGGGAATTTTANNGGCCTCCAAACGAATCNTGGGAAAACCC  
CCTTTGGGGGGGAAANAAACCTGGGTTACCNNGNAAAANGGGCCGNNNAAGNAAACCCNNGGGC  
GGNTTTTNTTTTAAANNATTTTNTCCCNNGGGCGGCCCCCCTTTGNGGGGAAATTTCCCCCAC  
ATGNGGNGGCTTTTNTGGGGGANCNANCTNGGNNNAAGTGGGGGAAAAANGGGGGANAAA  
AGNTTNTCCCGGGGAAAAATTTT

SEQ ID NO: 210 ACTCGGGGAAACGGAAGTGNNCNGGCGGGTCTNACTGACGGTNACGGGGC  
ATATAGGCTGTTCCNANAGCTGCGGATGATGAATGCCANNAGGACTTGGATCTGAGCTAAAGGAC  
AGTATNCCANTTACTGAACCTTTCACCAAGTGGACCTTNTGAAAGCATGATCTTCTCNGAAAGGTT  
TTTTCTTGTGGTNAAAAATGAACNTTNGCTTANCATCCCCTAGAATTTATCAAGAAAAAATTT  
CCANCTCAACCAAGATAANATTGAA

SEQ ID NO: 211 ACAAATGAAATTTAGGACCAGAGAAAAATGCAAATTAACCTGAAAGTTTAA  
GACAGGGAGAAAGTTAGAATGCAATAGCATAAATATGTTCTACCCAGATATTATATTAA  
AATGGCTAATTTTATTGACTTTTCTGGTAGAAAAACAAAGGAGGTAAAGCTATCTATGTAGTGATAT  
CTCAGCTAGTGCATGTGGAAATGTGTGTGGGCAGTTTGGGTGGTCACAATGACTGAATGCCTAGCT  
GGCATTAAATGTCTGGAAGCCAGGGATTCCAAATGGCTATCCTGGACAGGGGACTGGGTTGAGGGG  
GCCAATGGGGAGGCACTCCCATATGCAAAGAATTGTTCTGCCCAAATGCCATAACACCCTGCTG  
AGAAAGGCTGAGTGAAATGTTTGTCTTAAGTAAAAAANGGTT

SEQ ID NO: 212 ACGCGGGCAGGGGTAGAATGGAAGGAGAGCGGCTGGAGAGGACAGGTGGT  
GGAGGGCCTTGGCTTCTGCTAAGTGAGATGGGAACCACTGGAGGGTTTGAACAGAGGAGTGCCTT  
GATTGATTTATATTTTGAAGGGTCATTCTAGCTGCCATATTGTGAAAACTTTAGTGGACAAGGG  
CAGAAGGAAGAGGGAAGACCTGTTAGGAAGCTACTGCAAGGTTCCAGGCTTGGGCCTGGGCCAC  
AGCAACAGCAGTGGTCAAATATCTAGATTTATTTTGAAGAGCCAATAGGATTTGCTGAGAGTTT  
GAATGTGGAGTGTAAGAGAAGGAAGAGTTAATGATGACATTAAGGTTTGTGGCCTGAATAGCAGG  
AAAGATGGAGTTACCAGTTACTGAAATAGGGAAGGATGGGCTGGGTAAGTAAGGAATTTGGTGCA  
AAGCAGCTGTCTGTGGTTGGAATGGGAGGTTCTGCTTGCAAATCAAAGTGGAGAGTTCTCTCAGGT  
CAGGTCTGCANCAAAGCTCGAGACAGGGATCTGAATGCACTTGGTTTATTGTTGGGGGTGCTCTCA  
NAAGGAACCTGTGAAAGCCTTATCAGTCATTTATTGCTGTGANAAGTTCTCTTGGAAATGGGGTA  
CCTCGCCCCGACCACCCTAA

SEQ ID NO: 213 ACTTTTAGGAGAGATGGGATTTACCATGTTGGCTAGGATGGTCTCGATCTCT  
TAACCTCGTGATCCGCCCACCTCAGCCTCCCAAAGTGCTGGTATTACAGGCATGTGCCACTGCTCT  
CGGCCAATTAATTTTTTTTTTATGGAGATATGGGGTCCCTTTCTTTGCCTAACTACTGGGCCTCAA  
GCCATCTTCCACCTTGACCTCCTAAAATACTGAGATTACAGGTATGAGTCACTGTGTCTGACCAT  
CATGGAAATTAACGACATGCTTCTGAATGACCAAAGAGTTATGAAGAANTTAAGAAGAAAAATA  
AANCAAATTTCAAGACATGAAAAAGAACNACACCATAACCAAACCCCTGTAATACTGAAAAACC

ATGTAAGGGGGGAAGTCATAAGATTAATGGCCTACTTCAAATTAGAATTTTCAAATAAAGAAGCT  
TCTTTNCTTGTCAGNAACTNGATTAAAAGGAGNAAACCAAATCAAATTGTTATAGGAAAGAAGAA  
TTGNGATCAACCGAATTAACCAAATTAAGACTTAAAAATCCACNGACCATNGAACCGAAAGTGGT  
TTTTGTAAANATAAAAAAAGTTGAAAACCNCTCTGTTNCTAAGGAAAAAGANGGAAGACCC

SEQ ID NO: 214 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTAATTAAAGTATATAATGTAAA  
TATGGTATTGAAAGTATAAAAAATTAAAGGCTTCTGAAAACTCAAGAAGGGCAAAAGGCTATACA  
AGCAAGTAGTATACAGATTGTTTCCTTAAAAGATAAAATTTATATCCAGAANTTAAATCCCAGATG  
ATTTTTTTTNTTTTCACTTTCCCAAACCTTGGGAGCNTGAATGCCTCTAAATNTNNCCAATGAANTTT  
ACATTTT

SEQ ID NO: 215 ACNCGGGGTCTTTCCCATCTTGCAAGATGGCGGGTGAAAAAGTTGAGAAGCC  
AGATACTTAAATAGAAGANACNCGNACCATAGNANGTTGNTNCTTNGTANGCATGGGAAATAATT  
NTAACCTTAAAT

SEQ ID NO: 216 ACCTGTAGTCCCATCTACTAGGGTAGCTAAGGCAGGAGGATCGCTTGAGCCC  
ANAAGGTTCAAGGCTGCAGTGAGCTATGATCATGCCGCTGTAATCCAGCCTANGTGACACAGTGAG  
ACCATGTCTCTAAAAAACTAAAAAATATTTTTAAAAAATTTTAAATAGACAATACCTAAAAACT  
ACCTTTAAATATGCTATGGGGCCCCGGATGCATTGGCTCATGCCTGTAATCCCANCACTTTGGGAG  
GCCCANGTAAGAAGGATCGCTTGAGGCTAANAAAGTTCAAANAACAGCCTGAGTTGACAAGCAA  
TAGCTTATCTCTANAAAGCCTATTTAAANTNNAANAAATTNANNNTATTNATATTTTCTTGGGCA  
TGGTCAAGAGCCTNTATTTNCACTTNTCAAGAGGGCTGANGGGGGAGGATCACTTGNATCCAAG  
AAATCCAGCCTTTTAGGCTCATTGAACCTAGCTCACCCAACCTGNACCCCAACTTGGGCAACANA  
GTAAGATCCCATTCT

SEQ ID NO: 217 ACGCGGGGAGTTCCAAGTAGGTAATCCTTCTGAGAAGTCCCACCTTTCTGAG  
CAGCTGTGTTTGAAGAAAGCTAGTGGGAAAGTTCCAGGATTACATGTCAGGAAACTACAAGAGG  
TAANAAACATTTTGTGATTACAGTGTNTTAACNTTCTNCTGGGCTGAAAACCTTGCTTGTTT  
GTNAAAAATG

SEQ ID NO: 218 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGTTTTTTTTTTTTTTTTTTTTTT  
TCNCCANCCCCAGGGGTTAAATNTNTGAATCAAAANATCAGTTCAAANAGGACCCCTGNTTT  
GTCCTCATGCAGGGGTNAAAATNTNCAAAACCNCCCTGGGAATGTCCAAGCCCCAAAAANCCAGG  
GGCCANTCCCTGAGCAAGNGGAAAATTGGGTCTTGGAGTNNTAGGCTGCCTCCTCTTTTNCNNTC  
CTCCTAANTTTTATGANACTGNNGGGGNTTGGGGTAACAAACNNGNCAAAATAATTTTTTTGNGCT  
GGACCTCNTTCTGNAACTGGGGCTCAACTGGGANTTCTGGNTGAANTTGGCNGNCCNGGNGG  
TGGGTTT

SEQ ID NO: 219 ACAGGTTTTGCCAGTCTCCTATAGCATGGTATAGTGATAACTGATTTTTTAT  
AACAATGACTCANAGGCATTGAAGATCCATAACTATCTTCTGAATTATCACAGAAAGNAAGAAAG  
TTAGATNAGTTTATGTTAANTNTTAATAAATCATNTAACNATTGTTGTANTTGTGGNTCTT

SEQ ID NO: 220 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAAATGTTTATGGTTTTTATTTT  
TCAATTTTATTTTGGTTTTCTTACAAAGGTTGACATTTTCCATAACAGGTGTAAGAGTGTGAAAA  
AAAAATTCAAATTTTGGGGGAGCGAGGGGAAGGANTTAATGAACTGTATTGCACAATGCTCTG  
ATCAATCCTTCTTTTCTTCTTGCCCCACANTTTAAGCAAGTANATGTGCCNAAAAAATGGAAGG  
ATTCAACTTTTTCAGTTAAAAAAGAAGAANGAAGAAATTNGCCAAAGANAAAGTTTTTCAAATTT  
CTTTNTTTTTTAAATTTAAATGAAGTTCATTTTATTTGAAACAACTGGGCCNATTTCCCNCAANA  
ATTCCCTGGTCAAGACCNCCGATNTCCAAAGGTGCCATTTAAANGAAGGGCAGGCNNGATNGCCT  
ATTTTTTGNATTCAAGATTGCTTTCCCCATCATTTGTCCTTTTAAAACCNCAATNTACAANAACG  
TGTAAGGTGNACCCTCTGTTGNCCTCACNAACAATNTTCNACANTCATTAGAACCCTGTAAAAATG  
ACACCCTTTTCAGGTTGCNAATCCCNCTCCATNTTNCATGCTGTTCTTGCATGGGGANG

SEQ ID NO: 221 ACACTCTATGTCTGCATTTGATTATTACCTTAAACAGACTTATTGGGTCACA  
GCATATAAATGTTTCCATCCTTGAAGGACACTGTGAGGTGGCTTTTAAAGATAAAGCTTCATTTGTT  
AAAAATTCTAAATGCATTTTAAATGACTATGCAAGAAAGTTCATTAAAAATATACTTTCTTTGGG  
GAGGCCGAGGCANGCCGATCACCTGAAGNCAGGAGTTTCNAGACCANGCCTGAACAACATTG  
GGAGAAACCCCCGTCTTCTACTNAAAAATACAAAAATTAGNCCGGGGCCGTAAAGTGGCANAATGCC  
CGGAATTCCAACNTCTANGGANGCCTNANCAAGGAAAAACTGCTTGAACCCTGGGANGCCNGAA  
GTTGCGGATAGCCCNATNTNANCTTT

SEQ ID NO: 222      ACGCGGGTAAGAATCCTTGAAATAAATGTTTTTTTAAAAATCTCTGGGAAGA  
GTAAACTTGATTATTTATTTCTAAGTGAAGTATAATTATCTTTAATTTTTGTGAAGTAGGGACTCTC  
CTAAGAAATTATCAAATGCAACTCAAATGCAATTTTTAAAAAGNAATCCANTTAAAGCTTTCTTTG  
GAAACAANTTTAAGTTTTATGGTTATGATTTGAATAAGCCNAAATTATCTATTTTTGCAAGAACCC  
AATTTTTCATGCTCTGCAAATTGNTTCTGTTTCTATGTAAAAATTTTAATTATNCTCCTANACNTT  
ATTATAAAAGATNCTGATTTGATTANCATTCCCTTNCCTGAANCACCNTNCTAAGACATNGAGAN  
TAAATCCTTTT

SEQ ID NO: 223      ACGCGGGACTCTTGAAATGGTTATCTTTGTGGATGATTTTTTTTTTAAAGCTGA  
AAGTTACCTCNTGAATAACTNGTTTAAAGTGGTGGGGGNTTAAATTTTCANTAGAATCAATTGAAN  
CAAAAATTTAAACCTACTCTTTTGATTTTCTNACTTTTTCANTCCTTNNCCATTAAAGCNCCCTTAAAAAT  
GTAAGTTANTTTTAAA

SEQ ID NO: 224      ACTATATTTGATTTTTAGTCTAGTAAATGTTAGTAACTTGTTAAATGCGTTCT  
AACAAGGGAAGGCAGGTATTGGGTAGAAAATACATGTTTTGACTGTATCAGCCATGGGTGACTTCT  
GCAAATTTACTATCTTCCAAAGAACGATACACTGTCTATTGTAGAGACTGCACAATCTGTCCATTAG  
AGCACAACCTCTTAATAGATTACCAGTTCTATCCCAATATACAATGATTTTGATGCTTCAAAAACA  
TTTTAAATACAGTATAAAATCACTTAATAATATTGCATATTAAGTCTATGGTTACATTTTTGTAT  
ATTAGAAGCCTCAGAGAAAGAGTTAAGAAGTATATTTTATGCAATACTATATGCCCTGCAATAGG  
TAAATACTGAAGTTAAATCTTTACTCTGATCACATTAAATATCCCAATATTTAAAATATTTAAAG  
CCATAAANGGAAGCTTCATATCCACAGTTTATTGCTCCCATGAGGTATCTGCAAGCCTACATAAC  
ATAATGATTCCTTGGTTTGGTTTCCAAACAAATGGATTTAATTTANTANTTCCNCCTTAAACCCCGG  
NAACAACATGACACTGGNTT

SEQ ID NO: 225      GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTANTATCTGCTTCTGCCTC  
CCCCGCCATNTTATGGNGGAAGCAGTCGCTATGATGATTATAGCAGCTCACGTGATGGATATGGT  
GGAAGTCGAAACAGTTACTCAAGCAGCCNAAAGTATCTCTACTCAAGTTGNGACAGGGTTGGCAA  
ACAAAAANAGGGCTTCCCCCTTCTGTAAAAAGGGGGT

SEQ ID NO: 226      GGTACTCAGAGGAATTTTTTTTGTGTTTGTGCTTTTAAAGAAAGGAAAGAA  
AGGATGAAAAAATAAACAGAAAACCTCAGCTCAGGCACAATTGTCACCAAGGAGTTAAAAGCTT  
CTTCTTCAATAGAGGAATTGTCTGGGGGTCTGGAGACTTACCATTGAGCCATGCAATCTGGGAA  
GCACAGGAATAAGTAGACACTTTGAAAATGGATTTGAATGTTCTCATCCCTTTGACAGCTTTTCTTT  
TTGGCTCTCTTATGTCCTTGGCTTGCTCCTCTATTCTACCTCTCTTTCTCCAGCAATAATATGCAAAT  
GAAGACATGTATCCATAAGAAGGAGTGCTCTTCATCAACTAATAGAGCACCTACCACAGTGTCAT  
ACCTGGTAGAGGTGAGCAATTCATATCAAAGGTTGCAAAGTGTTTGTAAATATATTCATGAGGCTG  
GAAGTAAGAAAGAATTTAAAAATTTGCCTAATTACAATGGAGAACCATTCTAGGNAGTGATCTTGG  
ACCCACATGAATAACTTTCTTGAAGGGCAACCCAAATCCATTTTATTTCTGNCTGGCTTGGCACCT  
NTGGAAGGTT

SEQ ID NO: 227      ACAAACAACCACTTCTCAGTAGAAAGTTAAGAATAACATTTAAAAACATATT  
CATGTTTTAGAGAACGAATGTGCCATCGTTGTATATTAATAAAAAATAAAGATTAAACCAGCTATA  
AGAACTACAATTACAAGTAGAGTGGCAGTGTTTTTAACTAATAAAAGTATACATGTTTATAAG  
TGCAGCATACCTGAAATCTTGATGTTTGTCAATACTTATGGTTGCTTCAAAGATAAATTTATGTGAT  
TATTTTTGAAAGATGTGTATTAATTTGAATAATACCCAGAAAAATTATAACTTAAAAATTCAGTT  
TTCAATATGAGAATCATTTATGTGTGTAATCACTCAACTAAGAAAGATCAAAGTGTTGGTATAATA  
TTACAAGAAAAAATATTCAAAATGGAAAGTCCATTTATGAATGTATTAATATTAATAATCCAAAGG  
TATGTTTTTTTATAATGNCTACATTATAATGNTTACAAANGCCATAAAATCATTTTCAGAAAGGTCT  
CATNCTNCAGATATGGCCCATAACTTCATTTCTANAAAAAAGAAGAAATGTTTTATTNTACCAA  
GAT

SEQ ID NO: 228      GCGTGGTCCGCGGCCGAGGTACTACTTCTCAAGGAGGATTCATGGTCTGTCCCT  
TTGCTCACTACAGATTTCTCCTCTTCTCTGGGAAAAAATGGTGAATGCTTCTGCTNCCTTNTACTNA  
ACTAGATCCCTTGATTATTACTACATTATTGGNTATCCCCNCGTTGACACCTNCTTAGAAACNCN  
ATCGGGNGATGTGTGTTNATTAGGCA

SEQ ID NO: 229      GCGAGCGGNCGCCCGGGCANGTACAGNGGCCCCCGTGAAAGACAGAATTG  
TGGTGAATCCTGGTTGTACGCCCCTCCAGTGTCANATAAGGGCTGCTGCTTGNGACGACACCGT  
GTCGGGGGGTCCCGNGGAGCTTACTATCCTAAT

SEQ ID NO: 230 CNCNGGACACTGCGCCATTTCTGTCCAAAGCTGGGCGAATCAGGGATNCCG  
 GTTCACAATGGATGCTGATAAAGAGAAAAGATTTGCACAAATTTCTTAAAAATGTGGATGAAATCT  
 CCAATTTAATTCATGAGATGAATTTCTNATGACCCAGTTGTGCAACATGAAAGCTGCCCTGGAGACA  
 GAAANGAGACTACNNGTTATGGAGGAAGACCATTGNAGGAGGATGAATGCATGACCACCTTGAAC  
 AANACTATTATCAGTCTTTCACACAACCTGNTCTNAAANAGTGCAGNAATGAAATAAACTCNGAGGC  
 CTTCTTGGNATCTGTGGAGAAGGATGCNAAGGANCTAGCCAGGTGAAGAAGGGAAAAACAAAGTC  
 TTGGCNGATGCCCTNNANNAAAAAAGGGAATGAACCATTTGGTGAANGGGAATNATTNAAACAAG  
 CTATCCTGCGCTCACTGAGGGTTTNGAGANTCTNNACGACATGANNGGNTCTGTTNCTCNGCCNT  
 CNCNCNNAGGGTTAATTCATNTCACTTCTTTCTTACANGTGGGACCGANTTNCGTNCCAACCT

SEQ ID NO: 231 TAAAGATCCAGCGTTTCCCCTGGAACCTTCTCGTGCCTTNTCTGTTCCAACCTT  
 GCCGNTACCGGATCCTGTCCGCNTTNTNCCTTNGGAACGNGGNGCNTTTTTATANTTACC

SEQ ID NO: 232 TNNAGAGAGAAAGACAGGGCTCTGAAAATACTGCCATAGGCTCAAGTTCCAA  
 AGTGCTGGAGTTACAAAAGTATAAAGGACAACGCAAAAGGACTTTTTAGCCAAAGAAGAACCAGA  
 AAGGAAGAGTCTAGGGAATTGGAGACGTCGCGAAGGAATGTTGTAAGGAGAATTCAAGCCTAA  
 AACGGTTCTGAGCACTGATTTTCATCAAGCCGATGGAGGCAACAACATATTCCTTAGGGAAAATTA  
 AGATGCACAGTATGCTTCAAGATGATGGCCAAGAACAACCAACTCAACGAGCCTTCAGGAATTCA  
 TGATTTGAATATCATGTTCAAGGCTGCCTTCTAGATAAAGATCAAACCAAGAAAGTCCATGTGTCA  
 CTGCTCCGTGAAGGCTCCGGGTGGAACCTCCCTGCCCCACTCCACCTTTAAATTTNCAGCCAAGTG  
 TCTTACCTTTCTCAGCGCCCTGGTTACATGTCCATCTCTCTNTTTAGAAGTACAGACATCAGCAA  
 CTCGAGAAAAGAAGGTTGNTCTGGGAGAATCTAATGTGACCGTTATTGNAATATATNTCTTTAC  
 TCAAGTTTTTGGCTTNGGGCTGCTAAATATAACCTTTTNACTGGTTGNCTTGAGNGGN

SEQ ID NO: 233 GGTACAACCTTCAAACATTCCAGTTTTTTATAAAAAAAGGGGCACACAATCGT  
 GGTTTTGATCCCCTTTTGTGTTTTTGGACAAATGTTTCTACAAATACAGATTTCAGCAAACCCAAAGGC  
 TGCAAATTTACTTGACAGTAATCTGCCAGCCATTTACTCCACTTAATCCAGCTGAACTTGAAAT  
 GCAAACCAAATGTCCATGGTCATTAGCAATCATAGCAGGTAGAAAGGCTTTATAAGTCCATAAAT  
 GTGCTTTGAAATTCACATCAAATGACTTTTCCATAAGCTCATCTGGACAGTCAAGGAACCTTTTGC  
 CTGTTACGATTCCGGCATTGTTGATTAGGATGGAACACTCGCCGACTCTTTTTTAACTTGCTGGC  
 TACTCTATACACTCCTTCCCTTTTGGCTGCAATCGCAGGTATAGGCGTGCACTCTTGGTGGCTCCAGC  
 TTCCCGAACCATTCTACATGTTTCTCATTCCCTCCTTATTGATATCCAGAGAACAAGAACAGAT  
 CCAACCCGGCNAACTGCAAGGTAAACCTTTGAGTCACTTTCACCCCTGTGATGAAGACTATTTA  
 CCACAACGTCTTCCGNGNTTNGGAGTANGCAAAATATANCCTCN

SEQ ID NO: 234 GGTACGCGGGGGATGTGTCAGCTCCGCAGGGGTTTGGGGAAACGGCCGCTGA  
 GTGAGGCATCGGCTGTGTTTCTCACCGCGGTCTTTTCTCCACTCTTGGCTGGTTGGACCCCACTA  
 TGGAAAAGTTGGCCCTGAGCCAGAGCTCCAGCAGCCTTGTAGGGCGTGGCCTGGGCTTGGAAT  
 AAGTGGGATGTAAAACGAAGATCAGGAGCAGATTTGAAGAATTACAAAGTGAATTGGTGCCAGTC  
 AGCATGTCAGAGACAGACCACATAAGCCTCTACTTCTCTGATAAAAAATGTTGGGAAAAACACCTG  
 AATTAAGGAAGACTCATGCAACTTGGTTTCTGGCAATGAAAGCCACAAATTAATAAATGAGTCC  
 CAACTATTGNCATTAAACACTGATNAACTTTATTGTCAACCTANTGACATTATAATCGAATTGAA  
 CCAAGAAAATTATNTTCCAATCTGGGGGGNGGGANGGTTNTTGGGCCAAAACAAACCCAGGCTA  
 AAAANTTTTGAACAAATACTTANTTTTCTNAGGGGAAATTTNCTAACTTTTTCAAACCATNGANC  
 CGACCGAAAGTCCCCAATTTGGGGAATAAGGCTTTTTTTCCNANTANTNTGAAA

SEQ ID NO: 235 GGTACTCTGGACAAGGACAATCAGCATCTTCTCCCAAGGCAGCTCAGGGGCT  
 GGCATGAGGCAAAATACAACAGACTATCCCGAACTCCTTATTAGACCACAAAATACAAACAGCATT  
 TTCTTTTTTCTCTTTTTTTTTTTTTGTCAAGACAGTGCTCACTCTATCACCCAGGCTGGAGTGCAGA  
 GAGAGACTCCGTCTCAAAAAACAAAAACAAAAACAAAAACAAAAACAAACCTGTAGTT  
 TGGTAACAAATATTGAATCTTCTTTATTTTTAGTATCCACTCTCTTCCCACTGAAACTAAGAAACCC  
 AGAGCACTTCCACGAATTCTATTGAGATAAAGNGGAGGAAAATGCAACCCCCGCTACCTGCC  
 GGGCGGNGCGCTCGA

SEQ ID NO: 236 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGACTANGGTAAGTACTAGAGGTGT  
 GGGTGAGATGAAAATTNGGTAATAGGGACAAATGAAATAAGCCAAACTGTTTGGCCACNAAGACT  
 TCCANTCTTNAGGTTANTTCTGGTNTGTTAAAGNGGTTNTGCAGAGCTCAGNTCTGGGAATGAAT  
 CCCTTTTCAACTTCTTGATAAAGGCGTGACTTCCAAACTTTTCTGAGGTACANAAATCNTTTGGAG  
 CCATTTTGGAGACAAAAAGTAGAGGCTCTGNCAAGNCAATNCTGCATTGCATGCTTGGNCCACTGN  
 ATAANCCACGCCTGAGATACAAANGATGCACTACNCTTGACCCGCTTTATGTNCTCTTCTNTNCC  
 CTTNTTTNTNATNACTTTATTAGGGTAAACACCNCATACAGGCTTTTTCCANATGACTCCTATTTT  
 TGGGGGTGGGTAAAATTTTTGCCCCATAAACCAAACTTTGTGGCTATGCTNTTTGGCCCTTCNT

ACCTTTGGCNTTATTCCAAANTTTAAAGGNTNGGGGAAACTTTTCNGNTGCTGGNCCAAACCGGGG  
AATTTTCANTTAANATTTCTGAACCTTAAAGGAANATGGGCCNCAACTTNTTGGATCCTTTT

SEQ ID NO: 237 GGTACAAGTAAAGCCTGTGGTGGGATCAAGGAGCTATCAATATCAAGTTCAA  
GGATTTCCGTCTTACTTTTTGTCAGTCTGATATGTACCACCAAATCAAAAACATATCTGTGCTCTCT  
TTCTTAAAGAGTGTGAGGTCTCCAGTGATGATGTTAATAAAATTTTAAACATGGGTAAAGGAGGTAT  
CAAACACAAAAACCTAAACTTTGAAAATCTTAGGGAAACACTAAGAACTTTCCACAAGGAACT  
GGAAGGAAAGATCAAAAAGCAGCCTACACAGAATGGNCAGGAAGAGTTGNTCCTAAACAATGAGA  
TGAGTCTTCTCTGGAAAACACAAGTAAGGNGTGATTTTTATCATCAAGTTTCAGTGNAATTGACA  
TACTTAATAACTNCAGGGNTTAAATGGTACCCTNNCNTNGTATTTCTTTTTAAAAATNCATTTGG  
GGAAAAATGNGNTTATTGNTGCTTTTTGGAGGATGCCCTTGGCATTTTAAAGTTCTTATTTGGANC  
NTGAANGGTTTTATNANCNGACTTTTAAACNTTCCANTCCTTTCCCCCTTAAACGGGAATTAAC  
CCNTTTACCTTTGGNTNCCACGAATNCCTGATTTTTTATNAAAAAATCAGTNGGG

SEQ ID NO: 238 CGAGGTACCCTGAGCCAGGCGGGGACAAAACTGACAACTGCAGGATGTG  
GCCATACTGGGAGGAAAAGTCTCCACAAATGGCCTTGCCCCCGGACACCCCGCACACACACAC  
AAAACCCCTGCAGACCTACTTATACCCCTTAGCCTGTAAGCCCGGTGNCCTGCTCCTCAGATTTCT  
GTAGAGCAGCCCAGCAGGTAAATAAATTTGCTTGCCGACTTTGGGTCTTCTGTCTTTCTCTTGGC  
TAACCTTATAAGCCCATGTGCATAACTCCCTCTAGGTACGCGTCCCCAACCTATTTGGCATNAN  
GGACCANGNTTCGTGGAANGTAATTTTCCACAAAAGGTTGGCCNNGGGTTGGAAGGNGNGTG  
CCTNAAAAATTTGGGATATTGGNGGAAAAANGAANTNTTTGCCCNACNAANGGTTGNGANAAC  
GGTTTNTTGGATCTGNTGGGTGNGGNCCTNTNTGNTNACTACNNTCANAAAGTTNTNNNCACAA  
AAATTTCTTNAANAANANAAAGGNNCTGNAATTGGGGCNATGGTTGGAATGGAAGNGNTTTTGA  
GGNGGNTNTTNTACNTCTAAANTNGGNTTNCCTTTTTNATANNNGACAGAAACAGGNTGTTT  
TANNATATT

SEQ ID NO: 239 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGATGAAGTATTGCTCTGTGCGC  
CAGGCTGGAGTGCGGTGGCACAATGCAGCTTCCACCTCCAGGGTTCAAGCAATTCTCCTGCCTCA  
GTCTCCTGAGTAACTGGGATTACAGTCATGGGCCACCATGCCCGTTACAAGTCTTTATTANACT  
TGTATTGTGCAAATATTTTCTTTAGTCTTIGGCTTACCTTTTCATTCACTTAACAGTGTCTTTCAA  
GAGAAGAAAGTTTTAATCTACTAAAGTTTATTAGTTTTTCTCTTCTGGATTGNGCTTTGAGTGTTG  
TATCAAAAATCAAATCACAAACCCAAGGGCATGTAGATTTTCTCCTGGGNTTTCTTCTCAANNCT  
TTATAAGTTTGTCTTTACATTTAAAAATTATGAACCAGCTTGAGTTAATTTTGTGAAAAGGGTAAC  
GTCTGGGTCTGGGGTCAAATTTTGGACNTTTTATTGGTCCNACCCATGTTANTAAGGGGTAAAAACA  
CTTAAGGATCGCTTTATNTCCTGGGAGAAATATCCTTTTTTTTNTANGGAACACCCCTTTTNAACCC  
ATAANTTCTTACTTGAAAGCTNNTTAANCCNAAATAAAAAAGGTANATTNTTTNT

SEQ ID NO: 240 CGAGGTACAGCAACATGGCGGCGCCCATGGACTCTTAGAAAAAGGAGAAAGC  
TTTTTCTCTGTGGACTGGAAGGGGCATTTTTTCATGATCACTATTTAGATGGGTGCTGTTTTTCATGAG  
GAGAGTCTGGGAAGGCGGCGTCCGCTTTTTCTGACAAGGGAAGAGGCTACTTTGTCTTTTAAAGGAT  
TCAATGACTTCTGACTTGGAGGATGTGGACCTAGTGGCTAGACCCAAGGACCAAGCAAGAAGT  
CGTGGGGGGCCAGGAAGACAGGAGGATCACATTGGGATTCCAGACATAAGATCAGGTTTTAACC  
CCCTTTGGCCAAATTTTGGCTGAAAATGTTGAATTATCAACTCTGAAATTAAGAAAGTTTATA  
TAAAAACATTGCAATTTTCTTANAATTTCTGTATATATTAACATCATGAATGATAAATTCTCTTCA  
ATGTGCATGTCAAGGTTTGTACCTGCCCGGCGGCTCGAAAAGGCGAATTCA

SEQ ID NO: 241 CGAGGTACCTGGGGGCCAGAACGTAAGTTTTGACTCCTCTGCTAGGAGTGAG  
CTCAAAAATGGATATGATTCAAATACATAGATGCCTGTGGCCAATATTCGGATCTTCACAGTCCT  
CGGAATGCCCTGGCAGGGCTAAACTCCTTTTAGTCCAGTCCTCCTCAAGCTCAGACCTGCAAACTC  
TTCATTCTTACTGTGTATTGAGGGTTCCCTGAAGTGAAGGAAGAAAGTGTCTGGAGGGTGGGAGA  
GACCGTGTGTGGCAGAGTTAGAAACATCAGTCTATCTCAGGGTCCTAGACAAGTGATCTCCACAT  
AATCAGCAACAGATGGTCAGGCAGCACCTCCAAATATTTGTATCCTTATCAATCATATTTATGTGT  
GACTGCCAATCCATATGTCATGTGTATTAAGCTTTATTTTAGGTTTTTTTATTTTTTTTGGAGCG

SEQ ID NO: 242 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGNTTTTTTTTTTTTTTTTTTTTC  
CAAACCTGCAAAATTGCCCCAANTTTTATTTGTAGTCCNTACAAANGGGAAGAAAAANTTAAGGTTTTN  
TAACNCCACCTACTTGGGGANATGGGGAAATGGNACTGTCCCCCTCACCATCANCTAAACNTTA  
TNGGTCAGCAGGGACTTGNATACATACCAACTGACTGTCCCAANAGGANCTCAGTCT

SEQ ID NO: 243 ACTAGCATTTTCAATTACCAAAAACTATTAGCAGTAACATTGCCATTGGTAG  
TATTTTCAGTCAGTCCACTATCACTAAGCTAATGGGAAATCAATAGTAATAAAGTGGTAGATTATAA  
TGTAGCCATGAAAAATGTTGTTTAAATCAAGTTCCACAGTTCTTTCAGAAGAGAGGGAGATTGTTT

TGTATGTTCTTGATCTTTCTGAAGTTAAAGTTCATGTATTTTTAGGATGCTAAAAAAATATCCTGG  
ACNTTAAAAAGGTGCCAACTGTTACACACTCATTAAGAAAACACTGCTAGATGAATTATTAATA  
CCATGTGAATGCAGNTTATTANCCAANCAATGAATTTTCNGCATAANACAGTGTTATCTTTAANGCA  
CTGAAGTAAA

SEQ ID NO: 244 CGTACATNAAANACACGTCCACATCACANTTGCCCCCAAACCTGCCTGTGCTC  
CTCGATGGTGTCTCTCCCTNCATAAAACGCATGCTTATTGACCTTGGTTTTGATNTGCTTGGCCNTG  
TCNGTGANGATGAT

SEQ ID NO: 245 ACAGTCACACTGATGAAAGACAGGGAGGCCATGTGGAACCCCAGGGGCCAC  
AGAGCCACAGTGAAGCACCAGGCTAAGGCCATGAGTTAACACCACAAACCAGAAAGGCACCAGA  
CACTTGGGCTTTTGGTTAAGTAGGTTGGGCTTGCTCATTAAAGGACCTTTTCTTGAGACTCATTCTTT  
ATAATTTATTCAATTTACAAACACTTCAGTAGACTTCACAAACACTAATATAGTTCCTACCATATTA  
GTAACCTCTTTAATCCTAACAACTCTATGAGGGAAGGGAACTATTATTAGCCCCATTACAGAAAA  
GGAACCAGGCNCAGGGGAGATTAAACCATTTGCCCAAGATCAGGATGTCTGAAATGTTGATCTGGA  
TCAACNCTTCAAGGAATAATGACCTTTGAAAGATGCCCTTTTCTTAGAAAATCCTTANAATC  
CCTAGGAATTCTTAGGAGAGAATTTGTCCAGAAAATCTAGGACCTGCTGATAAAAAGGAAAAAGTA  
CACAGATTNTGNCCCTTGCCATACACTGC

SEQ ID NO: 246 ACTTNTTTTTTTTTTTTTTTTTTTTTTTTCCGACCCATGTGGACCAGGCTGGCC  
TNAAACTCNTGCCCTGGAACCCCCGCTCCNNGAGGGCCCNAGGGCAGGCNAACCGGCCTGAGCC  
ACANTGGCTCCCCCGCTTACCTNGGCCGANACCANNCTAANGGCGANTTCCTCACACTGGCGGGC  
GGTANCTANTG

SEQ ID NO: 247 GTACGCGGGGTTGAAAAATGGCGACTGTGGCAGAGTTGAAGGCTGTTTTAAA  
GGACACCTTGGAaaaaaaAGGGGGTATTAGGGCATTAAAAAGCAAGGATCCGAGCTGAAGTTTTCA  
ATGCCCTAAATGATNACCGATAACCCNTANCCCGTGGNCNCCATGGTAGGCACGGCAACTACCN  
CAAAAGTTGATAGGGCAAACTNTCAANTGGGTCGTCCCNCCCCCGCTACTTTTTTTTTNTTTTT  
TTTTTAGGGCTTTCAATNTTTTATTNAAATGCCNTGANCCAANATGGNTTTTAAATNTTGTNAAA  
AGCAGCCACATCCNTGGNCTGCNNATNNTATTTTAAAANCATNGATCNGCTCCTCCAACATATTG  
TTCCAACCTTTATCNTCTTNAACATACCCANTGTTTTTNAAGTTTCTNAATTCCGTATCCCNCTGNAA  
CTTNTTNAAAANACCCCCAACTACCCGTTTNTCTNNAATGCTTTCTNACCAC

SEQ ID NO: 248 ACGCAGGGGAATGGAATGGAATGGAATGCAATGGAATGGATTTCATCCGGAA  
TGGAATCGAATGGAATGGAATGCAAGCAATGGAATCAACTCGATTGCAATGGAATGGAATGGA  
ATGGAAGGAATACATTGGAATCAACCCGAGTGGAATGGAATGGAAGGACTGGAATGGAGTGG  
AATGGAATGGAATGCAATGGAATGGAATGGAATGGAATCAACTTGATTGGAATGGAATGGAATG  
GAATGGAATGGAATGGAATGGAATGGAATCAACCCGACTGCAGGGGAATGGAATGGAATGGCAT  
GCAATGGAATGGATTCAACTGGAATGGAATGGAAGAATGGAATCAACACGAGTGGAATGGCAT  
GGATTGGAATGGAATGGAATGGAATCAACCCGAATACAGGGGAATGGAATGGAATGGAATGCAA  
TGGAATGGAATCATCCGTAATGGAATGGAAGGAATGGAATGGAATGGAATGGAATGGAATGGA  
ATGGAATGGAATAGAATCAACTCGATTGCAATCGAATGGAATGGAATGGAATTAACCCGAATAGA  
ATGGAATGGAATGGAATGGACCCGGACGGAATGGATGGAATGGAATGGAATGGAATGGACCGAA  
AGGAATGGAATGGAATGGAATGGAA

SEQ ID NO: 249 ACGCGGGGGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCGT  
GGGGTCTGCTGTTTAGTCGCTTTCAGGGTCTTGAGCCCCCTTACGACCGTCACCATGGAAGTG  
TCACCATGTCAGCCTGTAAATGAAAATATGCAAGTCAGCANAAAAAAAAAAAAAAAAAAAAAN  
GTT

SEQ ID NO: 250 ACTCAGGGGAGGCCAGGANGGCCTTGANCTTGGGCCGGGCACTGAGGCGCC  
CCACATATGCTGAGAGCAGGGGGAACGCATCCAGNCTGCCANGGGCTAGGACCTCNTGGATCANC  
ANCNANTNCAGCAGGTTGTATTACGATAAGGATATNTGGTTTCCACNATTNAAGGINTTGCCTCA  
CCTGNTCTGGGACACAGNGGTCTAAAAAGGCTTAATATTNCCCGGACAGGGNCCTTCACATANTC  
ATTNCTTGGCCACCTCTTTNTTT

SEQ ID NO: 251 GAAACANATTAACCACATNCTCCTTCTTTGGGCTAGCAAGGTCCAGGGCTGC  
CTGGAGTCTGNCTCTACCATCAGGATANAGNAATCTTCCTGCTNGGATATAAAGGAGANGACNNA  
CATGTGTTTTTCACTTTNAGAACCATGTGGCCCGNTACCTATTGTGGATAATAACCCNNGTAAGA  
TGCTTGTGGAATGNANAAAAGTGCTGGGCATGAATGGTACTACNTGNGGCNGCTNTCTCAACTC  
AGCCAATTATGAGGTNTGTNATTGCCCATTGNAAGATGCTNANTGGNTAAACAANGTCTGGCCTT

NACCTTNCAAGANGGCNTANCCCCATTGNAAGAGCTCCCAN TGCCAAAATATA

SEQ ID NO: 252 ACTTTTTTTTTTNTTTTTTTNTTGCNAGATGTTCTTTTCTTTCTTTTCC  
TTTTNCGNANANACAGGGTTTCNCCNTGCTGCCCAGGCTGGTCTCAAAGTCCTGGCCTTAAGCNAT  
CCACCCACCTTGGCCTCCCAAANNGCTGGGATTACNGGCATGAGCCACCNTCTCTGACCANAGTN  
ACTTTTGNAANGGGCACTAATNCCACACNTGAGGCTCCACCTTNACNACCTTTTCATTTCCCAAAG  
GCTNAACNTCCTGCCCCNTCCCAATGAGNCCAGTGGCTTAACCTTTTCTNTNATTNTTTTNGNAA  
ATNATTGCTCATTTNGGGNAAAACCTCCAAATGCTCCGGTNC

SEQ ID NO: 253 ACCCAACAGGAGGTTTTTCAACTTTTGCTCTTCTCCCTTCTCCACCTCTCA  
CATCCCCAGAGCAAATTCAATGTAANTATACANTTTCTTCTTTCTTTTACAAAATATTNTGTTAAN  
ACTTACATGAGGTCCTTGAAAATTGGTCTAATATTTTCGCTTNTAAAANCTAGATACANAGCCGGGG  
TGTGGGTGGCTCACGCTTTGNAATCACAGCACTTTTGGGAGGCCAANACTGGTGGATCACNAGGT  
CAGGAGATCGANACCATNCTGCCTAATACCGNGAAANACTGTCTATACTAATAATTCAAAAATA  
TTANC

SEQ ID NO: 254 AGCGGCGAGGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAAAAANAAAA  
AGGAGCATTAACCTTGACTATGTCTTTANCTNCAGCCACCTTTTAAAGANTAAATTGCTGGGCAGG  
NGGGGGAGGGCTANTCANGNAACGAAACTGTAAGCCGGACNATNTGTGAGGAGGGGAGGTTAT

SEQ ID NO: 255 ACTTACATGTGTGAACACATATAAAGTGTGAGGTTTACAGACCCTGGCTCAA  
GGACAGTCTANGATGGGAAAGGAGGTANGGCGAGAAGAATCACATATTANACTCCCNNGGTGCTT  
NAGCCTCACCTATNCAAGGGACATGACNTATGGGGTNTNNTTANTCCATNCCCAGGTNCTATNCT  
TCANGACTTGAAGTTCCTAATTNGTATGNNGGAACCNCAANANCACGTTAACTGGCGGACTCCTG  
GNT

SEQ ID NO: 256 ACGCGGGGAGGGCCCCAGCCATCTCAGGCTACNCTATCCCAGGATCAGCATGG  
CCGCCTCCAGTGGATAATCNCCTGGCCTTGGCTGNCCTCCTTGTGTTGTGGACANGGNAGNGCCTT  
AGGCNTCANGAAAGCTCCCTTTNTCATGAATGCCANTNNGTGANCAC

SEQ ID NO: 257 CGCGGGTTTGAAAAGTCTTTGGCAATGANATTAAACTANAGAAACCAAAGGA  
ANAGACAGTNANANAGANCGAGATGCGANAACACTTTTGGCTAAAAATCTCCCTTTCAAAGTCAC  
TCGCGATGAATCGAAAGAAGTGTGTAAGATGCTGCGGAGATCANATTAGCANNAAGGATGGNA  
AAAGTNATAGGGATTGCAT

SEQ ID NO: 258 TCGCGGCGAGGTACACCAAGCTTCATTTTTGTTTTTGCNGGCTGAAGTCATG  
GCATGCAATTTTTGCATTTACGATTCTCTTGGGCATGCCCTGTGATCCCACCNCNCNGATTGAGTG  
CNTAGCCANTTGTGATCNTACTNTCCANATTGACTTCTTCNTGGNCTTTCCNAAATTTTACAGA  
GTTGACTG

SEQ ID NO: 259 ACGCCNTTCCGGCCAACANATGATATGCAAACCATTGTTGCTGTGGCCGAA  
CAAANTAACAAAACCTCCTACCGACACTGCAACATTGAAACNNTGGANGTGCCNATAAGGCTATT  
TTCTGACCTCACCTANTNTNTGCAANAGTGCANAANCCCCACAATNTCTNTGCTGGGCATCTCAC  
TGGAANATGGGGAANTGGTNGNCCCTGATTCCACANCCACCCATNTNATGTCCTGGGTTTNTGAG  
CNAAGGATCAAGCTNNGNCTTAGNGGCCTNTTAANCAGCCTGTTTCAAANTGCCATATGGACNAG  
NTGCATATGGGGAACATTGGGACTCTNAG

SEQ ID NO: 260 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTTTTTTTTTTTATTTGNNAATNG  
TTTATTTATCAANAGGAACATTTINTIANCCACATATTCATGTGTCATAGTTCAGGAACNCAAGT  
CAGGGACAACTTCTAGGNAATTCAACCCGAAAAAATTNTTATNTTCCAANATTNCTTTACNCTN  
TGAAANATCCAGCCTTCTNATNTCCTNAAAATCTTTNATGACNTCGGTANTTTCTGAAAAAAT

SEQ ID NO: 261 CGCGGCGAGGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGANAGAAGGANCCA  
ATGCAATGGCTANTCTTTACTGGAAANAAGAACTTTACAGCACATTTTGAGTAATTNAGNGGCA  
GGTGCTCANTGTGCCAGCTGNTGGGGCACCATCTTGTAATGGGCTCCACANAGGGGGGCATCNC  
GGCCTCGCCTTTGNGCACCAAAACCATAACAACGCTGGTTTTGTCTCTTANAGAT

SEQ ID NO: 262 ACCCTGGCATTGCTGACAGGATGCAGAAGGAGATCACAGCCCTGGTCCCCAG  
CACCATGAAGATCAAGATTATTGCTCCCCAGAGCGGANGTACTTTATTTTTTTTTTTATTTTTNCA  
AGGGTATAANCATTTAATTNAATTGANGGTAGNACCAATNCAAANTANGTTTGGNTNTTTATAA  
GTTANAGACAATGAANNNTCCNCCTTATGTTTAAAAATTTTAAAAATATCTNACATCNTTCTGCCA



GATGGGTTTTATTAAAAAGNGGAATTTTNCCTTGACACCNTCTTTTGGGAATTCANNTTTTTGAATGCT  
 NGATTACCCNATACCTNTTNTAAATNNTCTNTTTGTTTTAAAGGNTTTTCCAAAAAAATTTGNAT  
 AAATACG

SEQ ID NO: 263 ACNCGGGATGTGGAACATCTTTGCATGTTCTGTCCAGCAAGTTCTCTCTCCA  
 TGAAGACCATAACCACTAANGTGACCCCATCTCAGTCTGTCTGTTGGTAGACCTCCTGCNATCAAC  
 TGAAGGCCTGAATNANACAAAAAAGTGGCTTCTCCCATNCAANAANTGAANATTTGACATATCC  
 CCAAGACTGGAAGTGAAGTGAAGTGCATTAGGAAGCTGTATGTTTTGNGGAATTTAAATTTCAAATA  
 NAAATTGAAANGTNNCTAGANTNTTAGGACGTTATACCGTNTGAAC

SEQ ID NO: 264 ACAACCCTTCTCACCCCTGTGGGTTGGAGCCGAGTCAGGCCACTATGGGGAAG  
 CAGTTGCCCCACAAAATGTGGGTTTGCTGACCTATTCTAACTGTTGAATATGCTGCCCCATTGCTG  
 AATGAAAAAATGACTTTGGGGAACCAAACTNGCCCTTTGCCCAACTTGCAACCCCTTTNCCNNC  
 CNTTNTNNTNTTTGNTNGTNGTACNACCATTGTGGAAAAAATGNNAATTTGGTCTTCCAACCT  
 TAATGCNATGNGAACNATNAANTGGTNAGAAAACCAAACTGNTTAAAACCCCTTTCTCTNNAACN  
 ATTTATTTNTNATTTGTAAAAACACATATTAACNAATTCNTTTGCCTNAGGTTCTATTAAATATTTG  
 ANTTTTTTTTTTCTTATTAANAAAAATNAAATCCCTTACACCCNATATTTNNTTTATTCTNAACCGAAAA  
 CCCTTTNGGCNTTCCNAATTGAAANGGGGNTTCTTCTTAACTTTCTTTTTTTCTTTTTTGGGTTAN  
 AANGCCNTCATCCNGAAAGGGGCTNTGGGCTGTTAAAAACNNGGCTGGCCTTNGTNTAAANGGAG  
 CCAA

SEQ ID NO: 265 ACTGTTAAAAATGTTTCCATTGTTTATTCATCCACTGGCATTAGGTATACTTCT  
 GGTTCATGAATGAGCTTTTATCATAAAGTGAAGGCTAATTTTTGTATTACGTATCAGGGGTTGGAAC  
 CACTGTCTTCACTCCATACCTACTCCCCATTGGCAGTTTTCATGCAATGTTTTCTCTCAAGTTCA  
 GACCACTGTACTTTNTTTTTNTTTTTNTTTTTCGGGGACATTTCCACATGCTTTATCCNGCAATC  
 AAAATAATNAAAANCCATCTCAAATTATTATNCACNTACAAAAATAGGGTACCTNGGNCGGNACC  
 CCCCTNAGGGGNAAAATTCNCCA

SEQ ID NO: 266 ACGAAGAAGTCCTGGCAAAAATCAGCTCCACATCCACAGATCGGCTCACAGT  
 TCTCAAGACCAAGCCACAGTCTATACAAAGGGATATCATTACTGTCTGCAACCGACCCCTTACACGT  
 TGGCCAGCAGCTGACTCATATAGAGCTGGAGAGGCTCAATTATATTGGGCCAGAGAATTTGTT  
 CAGGCGTTCGTGCAGAANGACCCTTTGGATAATGACAAGAGTTGCTACAGTGAACGGAAGAAAAAC  
 ACCGAAACTTAGAAGCTTACGTTGGAATGGTTTAAATCGCCTCAGCTCTTGGTTGCTACAGAAATCT  
 GTATGCCTGTTAAGAAAAAACACCGAGCAAGAATGATTGAGTATTTTCATTGACGTAGCTCGGGA  
 GTGGTTTTACATTGGCAACTNCAACTCCNTGATGGCGATATCTCTGGTATGAATATGAGCCCANCT  
 NTCNACTAAAAAAACTTGGGCCNAAAGTGAAGACTGNAAAAATTTGACATTCCTTGACNTNCAG  
 ATGGACCCTNCAANCAATTTCTTTAATTTTCGAAACACTCTTCGTGGGGNAAACNAAANGGGTCNT  
 TTACCTGGTCATANTAGTAAAAAANGATTGGNNATNCCNTTCTTNAATNTTTAATTCAAAAA  
 TATTTNTTTCCTNAAT

SEQ ID NO: 267 GTACGCGGGGGGGATACGCCGCGNGCGCACGGCANTTAGTGGGTAGGCCTGA  
 ATAGCCGAGGAAAACTGAGCCGTGGGCCTCANAAAGAAAGTTAANGCACCCGCAAGCCGGGCAAC  
 TGCCCTCCTTCCGCGCCGGCGGAGCGATTNAAGTGAAGAAACAATGGCCAGCAATCACAAATCTT  
 CAGCTCTCGCCCTGTTTCAAGAGGTGGAGTTGGGTTAACANGAAGGCCTCCTTCTGGGATACAACC  
 CCTATCAAGGAAAAATTCNAGTTGGCAACTGCNATTGCCACCTTGGGACAGGCAAGACCAAGGT  
 TCTCNTGGGTTNGNCCATAGGGGACTGGTGGGANTTCTGCCTTCTNAAAAATNNAAGNTGCCCC  
 NTCGNCCTGAAACACAACAANNGTTTTNTNTGGAATTGAAANCTTGGGACNATAAGGTCCCCATGA  
 GGCANAATTTTAGANAAATCTTTACTTATCTNNGGGCTTTTTTAAAGTAAAAA

SEQ ID NO: 268 ACGCGGGGCTATTGCCTAAGGACTGCTTCCCCTCTTCAACAGTGAAGCTGCA  
 GGCCAACCACATGGAAAGGAAAAAGAGACATGAAGGGGAAGCAAATGTGTGTTCCAATACTTGA  
 AGCACAGTTTATGCTTTTCACTTTGGCATATATCCTATCTGCCAAACCCTATGTATATGCCACACA  
 TAAATTTCAAGGAAGCTGCAAATGTGTTCTTTAACTTGACTAACATGTACATTTGTCAAAAAACAA  
 AATTAGCCAGGCGTGGTGGCGGGCATCTGTAGTCCCAGCTACTCGGGAGGCCGAGGCAGGAGAAT  
 GCGTGAACCTCGAAGGTGGAGCTTACAGTAAGCCGAGATCACGCCACTGCTCTTTAACTTGGGTG  
 ACAGAGCAAGAATCCATCTCAAAAAAANGNNAAAAANNCANTCCTAAAAAATTTGTACA  
 TGGCATTGTTTTACAAAATTTTCTAANNGGANGAGNGAATATTTNATTTTCTGATGGATCCTTAC  
 TTTATTAANTGANGAAGGCAAAGNAAGTATGCNATATTTTAAAAAANAANCAGA

SEQ ID NO: 269 ACTTTTTTTTTTTTTTTTTTTTTTGGGATGTTNGGNGGTTTTAACTTTGTTATGTC  
 CATAAGGTCCTTCCAGAATCTACATGACCTGTCTCCTTACTGATCATTTCTTGTCACTCCTTTCTCC  
 TATACTCTATCCAGGCCATACTGACCTTCTTGCAACTTCCCAAATAAGCCTAGTGTGTTTGGCTGG



AGGACTTTACTTTNCAGCTCCCCCTGCTGGAATGCTTCCCNCCCCAAGTTGTACCTNGAAAAATGG  
GAANAGTACGGNACCACAGTTTATTACCAATTCTGACGGAAGTCCNCCTTC

SEQ ID NO: 270 ACTTTTTTTTTTTTTTTTTTTTTTTTTNTCCCAAAATGTGTTTNTTGANATGGTTT  
CCCCTCATCTTGACNCCANANTGCTTTTAGTGCTTGCTTCCCTCCTGAAGGAACATCCTTCTTGTAAG  
CCTTGCTTTTCCNCCTGTANGCTGNCNNANGACANTGGANCANCCAAACNCACAAAANTACCGNTT  
GTGCATGGNTAAAAACCNGGGTGATTTTATACCNTCCCTGGGCATTTACAT

SEQ ID NO: 271 ACTTGCACAGGAAGTGTTGGCCGCTTGTTGCATTCCGTTGCTGCTCCAAATTA  
AAAAAGTTGGTTATTTGGGACCTCAATCTCAACACAAGTGCTTNGTCCCACCCATGGGANCCCCG  
CGTTCCAAGAAAAANGGGTGTCCCGTGTGCTAAANAAAACTTGGGTGTGACTGGTTTCGATAAAC  
TCCAAAGCTTGTTCCTCATACACCCGTAATGGCTTTCTTTTTCTTTTATTGGCCNATCTTTGAAC  
CATTTCAACTTCANGANAATGGAAAAANGGCCTCTTTTTTGCCCTTTACTTGAAAAAGGGAAAGN  
TTATTCTGGGGGNGTNTGGGGGAAACNTTGTAAAAATTTTTGTNNAAAGGTGGGGGNAANCNTGA  
NATTNAANTAAANTGCNTNCCCNATATANTAAAAAGNGGGTNAAAGAANNNTGTCTTGCGCCT  
TNTTTTNTCTTAANAAGGGGNNTTAAAAATNATTCCNTGNANGNCTTCCAATTTNCCNTTAACCGN  
CCTNAAAGC

SEQ ID NO: 272 ACATTTGGCATGATCTGGGCCTATGCGGTCTTACAATCCCTGTATAAAACTAG  
ACAATGAAAAACAGAAAAACAAACAAACAAAAACAAAGAACGAGCACCTACCACATG  
CCAGCTACTGAGGCTATGAAGGTATTCTCCCGCCTTAGAAAGCCCAGGATTAATGCAGGATTGCG  
ACATTTAAACAGAACATTTCCATACAGCATGAGTATAAATGACTTTCCCAAGTTTACACTGAGAGT  
AACTGACACAGCAACCCAGCAAAGCTGAGCTGAGTCCCTGAATAATTGTATAAAAAAGGGGAGAG  
AAACAGAGTGAAGAAAGGGTTTCCAGACTCTGTCCAGGAAAGAAAATGAGCTCGTGGAGAGG  
AATAGACTTTCTCTATGAAAACAGAGGGAACAAAGAGGAAGATGTCTGGGAACCGAGGAGTAAT  
AGAGACCTGAGTTTACATCACTACTCTGCCCTCCCTANGTACAAAAGTGATACAAATACAATGGA  
AAAATGCATAGAAAAGCAGGAAAGATTTTGTCAACTAAAAACAAAATTATGTTGCTTCAACAG  
CAACTGGGATGATGCTTCGCACAANAGCTTGACAATCAAAAGAAAAAGCAATACTTAATATTCAA  
GCCAGATGTGA

SEQ ID NO: 273 ACATGAAGTCCTATACGGTATAATGAGGCAATCAGAGATATAAGGATTGGAA  
AGGAAGAAACAAAATTGTGATTATCTACAGATTATTTGTAGATTATGTGATTCTATANAAAACCC  
AAAGAGTAAACCTTCACATTATTAGACACAACATGGGAATTTAGCAAGGTAAGTGCATATAAGAT  
TAATATACAAGATTCATTGTGTTTTCTCTTTAAATATAGGTNNTAATGGCTTGCTGCTGACGAATA  
GAAACACACCATTTTAAATTCTCTTTCCAATCCCGATTAANCTAGACCACAGAGCCAAAAANTTCA  
TTGCAATTTAATATNTTCTTTCTTTTGTCTGGCTTTNNAGAGCANAATCTCCTAAAACCCANCTNTA  
AACTTCANCATGTGATATGTTGINTGATATCATNNTTAAAGACAAAAAGTTNAAACCCCA

SEQ ID NO: 274 ACCTTAGTGAGGCTCAAAAGGATTCTTTTGAGTCTATTTTACGCCTTATCTTTG  
AAATTCACCACTCGGGTGAGAAAGGTGACATTGTAGTCTTTCTGGCCTGTGAACAAGATATTGAG  
AAAGTCTGTGAAACTGTCTATCAAGGATCTAACCTAAACCCAGATCTTGGAGAACTGGTGGTTGTT  
CCTTTGTNTCCAAAAAANANNGTACTGCGGGGAAACGGAAGTGAGCGGC  
GGGGTCNACCTGACGGTANCNGGGCAAANAGGCTGTTNCGCAAANCTGCGGAAGATAAATGCCN  
NAGGACTTGATCTGAGCTAAAGGCAATTTNCCATTTACTGAACTTTCATCAAGCGGGACCNTTTG  
AAAGTNNNTNANTCTTGTTGCGAAATGGATNTTCTTGNTNNGAAAAATGAACTTGNGCCTTNTNNTC  
CCTTNATTNCCANNAAAAANATTTNCCNCNNATCCAANANTCAATGATATTTCCNCNCTAGAN  
AACATTNCGGGNNTTTTTTGCTCCANTAAATNACNTATGGAATTCANGGNCTANACCA

SEQ ID NO: 275 ACGCGNGGGATAACTACCGATTTGCACATACGAATGTTGAGTCTCTGGTCGA  
ACGAGTATGATGATAATGGACAGGGNATCNTCTNATTTCCGGCNTNNACATGTGACTANCANGTN  
TGAGGACNAGACTGGGGCTTAT

SEQ ID NO: 276 ACTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGCCATTCAATTAATATTTATNGAN  
AGGCTTTANCGNGCNCATNTCACTCAAGGNAAACAGTTGAAAANCTTATTATTTAAACAGCCCAA  
GACAAATACTTGAGTAATAGATAATATAATTACATATAAACAGGTACATGTNNAATCTCCANA  
GCAAACCTTCTTTCTCGAGTTTGCTGATGTCTGNACCTTCTAAANANAGAGATGGACATNTAAAC  
ACGGGCCGCTTGAAGGAGGTGAANACACTGTGCTGGAANTCTTAAAGGTGGANNGGGCATGGA  
GTTTCNCCCGGANCTTCCNGNCATNGACACATGGACTTTTCTNNGGTNGANCTTTNTNAATAAGG  
CAACCTTGAAAATGATNTNAAAAATNATTGAATNCCNTNAAACCTNTTAAAGTTGGTNNCCNTGGC  
CATCNNNTTGAANCNTACAGNACATNCTTATTTACCTNAGGAAT

SEQ ID NO: 277 GTACCATCNTCATGTTATTTAAAAAATNTAAGNGCANAAAGTAAAAAAACCAA  
GGAGCTCTGNAATAAAGTAACATAAAACACCACTGTAAGGATCACTTAAAAAACGATTCCCTCAC  
GGGAAATNNGGGCACTGAGTGTGTAAACATGTNAGCTGTACCTNCATAATGAGTAANCANTNTGG  
TAATGAANTATTTCTGGCCAAGCTTATTGCTCAAGCTTGCNTATNTATNTTTCTGATTCCGGCCAT  
TNCAATGTNTGATAANNATCAGACTTTGNNAAACTTCTGGGCTCTNGNNATTANNTGCCTNTTNAC  
TATAGCATGGCATNGNTTTTAAANNATNGAAACCCNAAGTGTTGATGTA

SEQ ID NO: 278 GGTACCTATATGGCTCCCCCAATTAATCACAATTCAGCAGCATCCACTCCCCA  
TATGTGGCTGAGATAACTTACTTTGACTGTCTATTTGGTTATCTCTCAGACAAAGCTTCAACGAAA  
TAAAATGGACAGGTTTGTGGACTGTGTAGAGCAGCGGAAAAAGTTTCCACATTTTCTTGGACTGCTT  
CACCTTCTGTGACGCGCTCCAAGGTGTCTGCTCCAGGCATGTCTACACACCACTCTCTTTCTGGTA  
AATCTTGT

SEQ ID NO: 279 GGTACGCGGGGAAGTAAGTGTGGTGTGGAAGCAGAGTAGAGAGAAAACTTG  
TTCCTCATTAGAGAGAGAGCCACACTTCTCACTGCTCACAATGAGAGGCCAAAAGATTACCTTGGGA  
CATCCAGATTTTCTATTGTGCCAGACCTGACGAAGAGCCTTTTGTGAAGATCATCACTGTTGGAAG  
AAGGCAAAGCCGTATGAAGAGCACATGCAGCTACTATGAAGACGAGGACGAAGAGGTGCTGCCT  
GTCCTACGGCCCCACAGCGTGTCTGAGAAATATGCACATCGAGCCACTGGCCCCGACGCTTCCT  
GCAAGGGTGC AAAGGGTATCCATGGAGACTGGCCTATTGCACGTTAGAGCACGGGACCAGCTTAA  
AAGACGCTCTACCCGGAATNNGGCATCACTAGACAGTCTGTCCTATTGGTCTNTCAAAGAATAT  
GGATAATCAGATTTTGGAGCATATGCCAACTCATCTTCAAGGTTCAGTGACCACATTTT

SEQ ID NO: 280 ACGCGGGGCCAGGAAGATAGGCAGCTCATCTGTGTCCTGTGTCCAGTCATTG  
GGGCTCACCAGGGCCACCAACTCTCCACCCTAGACGAAGCCTTTGAAGAATTAAGAAGCAAAGAC  
TCAGGTGGACTGAAGGCCGCTATGATCGAATTGGTGGAAAGGTTGAAGTTCAAGAAGCTCAGACC  
CTAAAGTAACTCGGGACCAAATGAAGATGTTTATACAGCAGGAATTTAAGAAAGTTCAGAAAGTG  
ATTGCTGATGAGGAGCAGAAGGCCCTTCATCTAGTGGACATCCAAGAGGCAATGGCCACAGCTCA  
TGTGACTGAGATACTGGCAGACATCCAATCCCATGGATAGGTTGATGACTCAGATGGCCCAAG  
CCAAGGAACAACCTTGATACCTCTAATGAATCAGCTGAGCCAAAGGCAGAGGGCGATGANGAAGG  
ACCCAGTGGTGGCAGTGAAGAAGAAGACACATGAAGGCTTGCT

SEQ ID NO: 281 ACGCGGGAACAGTCCCTTTCTATTGTCTATTCTCCTCCTCCTCAGTCTTTACT  
GGATGTTTTATATGAATGTATTGATACAATTTGGGGTCTTTTGTGACTGGATTCCATCACTTAGCAT  
ATTTTGTAGGTTCACTCATGTTACAGTGCATATCAGTATTTTCAATTTCTTTTATCGCTAAATAATATT  
CCATTTTGGATGTAGAGGACATTTTAAAAATTATTTCATTANCTGATGGGACATTTGGGTTGGTTCT  
ACTTTTGGCTATTATAAATAATGCTGCTATGAACATCTGTATAAAGGGTTNCTGNGTACTCCAAN  
CNANNTTACNCTTTTTCGGGGANTCTTTNGGTTTATGATTACNAAACACACCATGGCTTNTGTC  
ACAAATAGGGGACATTTAAANCCTTCCATACCAATTTNTTTTTTACCTNGGACANGGGCCCCCAT  
AAAAGCCTTCCANGCCACGTTTTGAATGTGANANACATTTGGCAAAAAAATTT

SEQ ID NO: 282 ACAATTTGTGCCATTAAACATTATCTTTTCATCACAACCCCTAGGTGAAGTAT  
GCTGGGGAAGCCTTGCTCTGTGTTTTTGAATTTGTAGGTGGCAGCCCATTAAGTGGTCATATGGTT  
CATGGCCAGCATTTAAAGAGAAACAGAATAAATTAGGAAATATTATTTAAAAACTTTTAGCTAG  
AGTTCTGTGTGGAACATGGGAATCATAATGATCCAAATAGATGGGAAACAGAAACATGGCCAGA  
AGAATATTAGGAAATGAGAGTAAAGTGATACTGACAGCCTTATTGCTCTGAAACTCAATATGAAA  
TGAGAAGCAAGTAGGAACTTTTACATTTTGATAGGTGATGAGCAGAAAGATATCTCAAGCAATTA  
ACACACATCCAGGTGAGAAGTTGGCTAACTTTCCATAAAGGGCCAGATAGTAAATAGTTTAAA  
TTCTGTGGGCCATAAGGTGTCTGCTGAAATTATTTAACTCAGAGAATGGGA

SEQ ID NO: 283 CTTTGCATTCTGATTCTTTGATATTTGAACTTGCCCATCTGTGGGTGCTACTCT  
CGTAAGACTAAAAGGCACCTACTAAATACGAACAAGGATAACCAGTTCTACCAAAAAACATTACCA  
ACCAAACCTCCCCTGTTCTTTTTCAGACTTTCCTTTTGCTTTCAACCTCTTAAACTGTAAAGGAGA  
GATTTAATGTTTCCCTATTATTCTGCTTTGACTCACATACTAAAATGACCACATGGCACTCCATTGA  
ATCTTTTCAGTTGTTTACGAATTAATAATGCCCATGTTTTACAGTTTTTTCAGTAGTCTCAGCAA  
TGGATCTGTAGTTTTCAGGACCTGTTTTTAAAGATCAAGATACTTTTATGTGTTTCAGAAAGTAGC  
TGTTTTTTCATTTTTTGTGGGTTTGGTTTTTTGTGTTGTTTGGNIGGGTGGGTTTTGGTTTTTTGCT  
TTTTAAAGTCAAGTGTNTGAGCCTGGTGTAAATGCCNNGAATATTATATNGGACTTAAAGTATC  
TGCATCTCGGTNAATGNGCCAAGTCCTTCATTAANGGCTNNCCCAGGGGGGTTTTTGAAAANCCTT  
TG

SEQ ID NO: 284 GGTNCGCGGGAGACGGTTTAAACAGAAACAGCGGCAGTGTAGTATGGCCAGG  
GATACCCATTCTCCAAGACTCACCATGCTCCTCTAGGTGGCTCTGGCCTTTGTTGACTCTCATACAC

AGACAGAAAAGAGATATATTAACCATGGGATGGGGTCACTCTGATTAGAGTAACCCATTTTGCTG  
GCCTCTTCCAGGGTCCCTGCCTGTCTGCACCCACTTGCAGTGCAACCACAGATGCACAATCAGGGT  
CCTTCCCAGTGGCAACTGCTATAGCTCCTTCACTGGCAGATCCTGCCTAACAATCAGAGAGCTTCT  
GCAGAAGGGGGCCCTGCCAGCATACACTCAATCACAAGCCTCCTNTCACC AATTTAAAAACACAC  
ACTTGGCTGGTGGTGGTGGTGGGTACGACTGTCACTCAACACTTTGGGAACCCAAAGGCGGTGG  
ATCACCTGAGGTGANGTTCGAGTCAATTNTGGCCAACATGGTGNAACCCCATNTTNTCTAAAAA  
AAATTTTAA AATTANCCGAGGNGGNGGCACATGCCCGTATCCTACTACTTNGGCACTTAGTCAG  
GANAAATCTTGATNC

SEQ ID NO: 285 GGTACAGTGGCGTGATCTTGGCTCACTGTAGCTTCTGCCTCCTGGGTCAAGC  
CATTCTCCTGTCAGCCTCCCGAGTAGCTGGGATTATAGTCACGTGCCACCACGCCCGGCCAACTTT  
TGTAGTTTTTAATTGAGATGGGGTTTTCTATGTTGGTCAGGCTGGTCTCAAAGTCCTGACCTCAGGT  
GATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTATAGGCATGAGCCACTGTGCCCGGCCTGCT  
ATTTTAATATTTGAGAGATAATGTTACCAACATGCTCATCCACAATAATTACCAAAATTCATATAA  
TTGTATTCAACCTGTTTTTACATATAAAGGGAGAAAGTGCTTAATTAACCTAAACAAAAAACCTACC  
TCAGTCTTAACCATATGACAGAATCTTTAAAAATTTAATGAGTCACTCATTTTCATGTCATATTTAGAT  
GAAACTAAGCCATGATTATTTAATTTAACCTNGCAAGGCTCTTGGACTNGGTANAAAAAAATTGGCT  
TAACCTTTNGAAAAACATANAANCCTTNTTTTAACANTTT

SEQ ID NO: 286 TGTACTGAGGAAGACACCATTTCCTTGACGGTGTCTAAGAAGCCAGGTGGATG  
TGTGTGGTGGCTCCAGTGGGTGTTTCTACTCTGCCAGTGAGAGGCAGCCCCCTAGAACTCTTCAG  
GCGTAATGAAAATCAGCTCAAATGAGATCAGGCCCCCCCAGGGTCCACCCACAGAGCACTACAG  
AGCCTCTGAAAGACCATAGCACCAAGCGAGCCCCTTCAGATTCCCCCACTGTCCATCGGAAGATG  
CTCCAGAGTGGCTAGAGGGCATCTAAGGGCTCCAGCATGGCATATCCATGCCACGGTGCTGTGT  
CCATGATCTGAGTGATAGCTGCACCTGCTGCCTGGGATTGCAGATCTAAGGTGGAGTGGAAAATGG  
TTCCAGGAAGACAGTTTCACCTCTAAAGGTCGCGAAAATGTTNCCTTTACCTCTGGAAAGTGGGAGTTG  
AAGGGGTACACACCAAAAAGGTATTTTTCCCTCACCAGTCTNAGGCTTNACTGGCTTTTTGAAAAA  
ATTCAGCACACCTTTNTTGNAACCTTANTTGTNANCNAGAAAANGGGCCNT

SEQ ID NO: 287      ACCGCGGGAAACTATATGCTATCTACAAGAAATTTACTTCACCTGTAAGGAC  
ACAAATAGACTGAAATTGACAGATGGAAAAAATATTCCATACAAATGGAAACCAAAAAGTAGCA  
GGAGTAACTATACTTATATCAAATAAAATGGACTTTAAGTAAAAAACTATAAAAAGATGGAAAA  
GACCACAGGTCAATAGAGAAAGAGGAAATAATAAATGTAAAAATACATGTATCCAATATTGATGC  
ACCTAAATACATGAAGCAAATGTTAATAGACCAAATGGAGAGCTAGAATGCAATGCAGTGATAGT  
ATGAGACTTCAGCATCCCACTGTTCTGCAATGGATTGATCATTACACAGAAANTCAACAAAGAA  
ATATAAGAATTTAAAAATGGACTCTAGACTAAATGGACCCAACAGGATTTTAAAAAAACATTNCTC  
CAACATCTGCAGGATCCGTGAAAATTAACAGATAGCTTGAAATACTGGATGGGGCAATCAATT  
AAATTTTAAAAATNAATTTAAAAATTTCCCT

SEQ ID NO: 288 GGTACAATAAGTGCCTTGCACATAAGAGTCCAATAAAATTTCTTGAATGATGATATGCTGATACATGTGTTCAATAATTTATTTACACTGAAGTCTACTAAATCCTAAACATCTAAACCA CAATTTAATNACATTTTGATGAAAATTATACATGGAATGCCTTATTTGAAAAAATACACATACATA TAATTCCTCATCAATACTTCGAGAGTTCATGGATTTTCTGAAGCTTACCTATGAATCCCAGAAGAA AAACCTCTAATAAGAAAATTGTATGGGGGGGAAAAAGTCTAACATNTTGCTCAATAAAGTGNCCTCT TGATTCTGCCCTCAACTTCAGAAAGCTCCTAAAAACAGTTCTGACTGANCCAAAAGTCTTCATGCNT AAAGGATCTGAAGATGAACTCNTGAGTNAAGAAACCGGCTTNCCATTNNGGCNAGNGAANGAGC ATAAATTAGANTGGCCCCAAACCCGCTCTCCATCCATTNTNCTAATNCTGGAGCTTCTTGGTNTCTAT CTTCCCTTAATTCAAAAGGTCTATNCTTGCACTTTNGCCTCTNGAACNNATTCTCCTTCATCTNAAGT NAATACTACTTTCNATTTANTTTNTTTCTAATAANCCGGTANGGCNTTNCTTNAAATCTATAC NCTANTNGNTNGTTA

SEQ ID NO: 289 GGTACCGAGTGCACCTATGTCTAATCATGTGTGCATGTGAGGAGGTGCTGGC  
TGACTGCATCAGCGGAACCCAGGCATAACGGCAATCTTTTGTTTTAGGTATAGTAGACGGAAATAT  
AAAATAATGGAAAGTTTTTGTCTGATTATAACAAAAATATTTACTCTTTTCATTAGCAAAACATTAT  
TTAAATATATCTATTTCTCCCTCTGAACATTAACTAGGAACACTGGGCAATTTTGAAGGTTTA  
AGTGCTGATTGTTAATAACAATAAATATGGCTGTGGAGAAGCACACAGCAGCAAGTAGCTTACAT  
ACNGTCCACGGATAAGAAATTTACTTCTCATCTGGGTTAGGTGTCCAGCCACCCAAGGNAAGG  
TTTGGCATTCCGTGTGCTTTTTTTTNCGGGTANGAATNAATTCGCTNGATTTTTCATTCTTNGGAA  
AGAAAAATNATTACGTNGATATCTTAANNTTTATTGACNTGATGCCAAAGGAGCAAGGGG  
GCAAAAANTGANCNTATGGTCCAAAATGCTNGNTTNGGATCNTGTCTTCAGCTNATNTTCCTTNTT  
GCAAGGNNNGCTCENNANGGCTGCCCTTNAANCNTGATTGAAGGNTTGGGGANCGGGGGGGGTN

TCNNTTGGCAAGCTTTCTGTTTNGNNAACCCACCCNGTTTATTCTATAA

SEQ ID NO: 290 CCAGCGTTTCCCCTGGAAACTCCTTGGGCGCTCTCTGTTCCACCCTTGNCGTT  
ACCGGATCCTGGCCGCTTTCTCCTTTNGGAACGTGGCGCTTTTATACTACNCTGAAGATTTTCANTC  
GGGNAGNCGTGTCACACTGGCTGGGGCACAACCCCGTNACCGACGTGNNCTTNTCCGGACTTTG  
GT

SEQ ID NO: 291 CCGGGCAGGACGCATACAATGACAAAGCCATTTTGGAGCAGAGGAACATGCT  
GACGGAAGAGCTCTATGGGAACACATTTTCAGCTTTACAAGTCAGCAGATCACCCAACCTCTGGACA  
AAGTGTTAGAGGTACC

SEQ ID NO: 292 ACTATAAGTAGATCCACGTATAGAGAGAAAATTGATTTTTGACCAATTTTCAGT  
GAAGGAAGGGAATTTTTTAAAAACAGATATTCCTAGAGCAACTGGATATCCATTTGGAGAAAACAC  
CTATTTTACTTTTATTAATATCAGATAAAATTGACTTTAAAGTAAAAATAATTATCAGAGGTAAAAAG  
TCACTGAAAGAAGAAATAAGAATTCATCAGGTAGAGACACCAATTCTATATTGGTATGCACCTAA  
TGTATGCCAATATATAGAAAGCAATAATATAAATGAATGTATATAATATTGNATACCAAGCTAAC  
TTTATATATAGAATTCAGAGAGAAATGATTTGGCTGAATCATAGAGGGAGAAAACCAATGGCTTA  
TTAGGCACATGACTTTGATCTGTGCCCCTACAGTTTGNGTGGATGGAGCCAAATCNGGTTTTTC  
AGATTNNGACCAACAATCCTTCANGATGGTTGGTNCATATTAACTGGTACCGGGATTGTGCCTN  
GATATTCTGGCCTGAATAATCTCTAGNNGAGNTATTNAANCTTCTCACTGGANNAATAANCTTC  
CCCAAAAATTNGNCNTNTGNCTTAAGGNTAGGATTTTTTCTGATGAATAAACCTNTTGGNGAAAG  
GNNTTGGATAACAANTTTTTNGCCTTGCGGGNGTG

SEQ ID NO: 293 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCNGAAATGAACAAATATTTA  
TTTATNTTTTATAACAAGTAAGGCANTGTTGCTTAAAGGAAGACAAACAAACATAAAAGATTCCG  
TTGACAATGCATTTTTTTCATCTGTTGCGGCACAATGCTTTTGTGATAATGGANATGNGACAGCAAAN  
TTCCAGGACATTCAGTNTTCGGCGGCAGNANTTAGGGCANATGACTGGCCGCTCAAATNTCTAT  
NTGTTTCAGGACAGTGGAAGGCTGNTTNTAAATGAGGCCAAAGCACNAGGTAGGTGGAAGGTT  
CTTGNTCGGGTTGAACCNCGACAGCGCNCNCAANAGACAACACTGAGGCAATGGGGAACAACAT  
TGCTNTTTTAANTGANNCCTTGGGTGTCNAGCGTGTGAGGGTTAAAAANGGNTTAANNCCCGCG  
TTNTCGGCCCGACNACTCTTAAAGGCGAATTTCAACACCTGGCGGGCTGTNACTAGGGGTTCCNACT  
TCGNANCCAACCTTGGNCTAAATCNTGGGCTAGCTGTTTCTTGGNGGAAATGGATTNCGTTANATTN  
CCNNAAAATNNAACCNAGAAANTNAATTGTTAAANCTGGGGGCCTAATNGAGGGGCTAACNANATA  
ATTGNGGTGNGCTCCTNGCCGTTTCAANNNGGAAAANTGTGG

SEQ ID NO: 294 ACCATCTCACTCAACTCTTGCAAGAACTCTAACGAGACTGGTATTATTATTCC  
TATCTTACAAAAGAGGAAACCGGCCAGGCGCAGTGGCTCAGCCTGTAATCCCAACACTTCGAGA  
GGCCGAGGTGGGTGGATCACCTGAGGTGAGGAGTTGGAGACCAGACTGGCCAACATGGTGAAAC  
CCCATTTCTACTAAAACAAAAATATACAAAATTAGCCAGCCATGGTGGTGTGGGCCTGTAGTCCCA  
GCTAGTTGGGAGGCTGAGCCAGGAGGGTCACTTGAACCCAGGAGACAGAGATTGCAGTGAGCCA  
AGATCACGTCACTGTACAGAAGTTTTTAAACCAAACTGAGGCATAAAGCAGAAAGAGCAAAAGA  
CACATGAATACCCTTCTTAACAATCTCTTCTACTTATGCCTCACCGAACCTTTGNACCCAACCCTCC  
AATCCCAACGCGCTNAAGTCAGACTGGGCTTNAGCTTTTGNCCAGGCACCCCAACATTTTTTT  
TTNNATTTGNGTGGCCACTTNANAAAAACTTCCNANAAANAATTTTTGCTGGCAGTTNCN

SEQ ID NO: 295 ACGCGGGGCTCCTGTCTTGTCTCAGCGGCTGCCAACAGATCATGAGCCATC  
AGCTCCTCTGGGGCAGCTATAGGACAACAGAACTCTCACCAAAGGACCAGACACAGTGGGCACC  
ATGGGACAGTGTGCGTCAGCCAACGCAGAGGATGCTCAGGAATTCANCGATGTGGAGAGGGCCAT  
TGAGACCCCTCATCAAGAACTTTCACCAGNACTTNTTTTTTTTTTTTTTTTTTNTNGAGGGGGGTCT  
TTTTCTTCTTGNTCTCAAAAAGGNCAAGGGAGCCCGACNAGGAATAAATANCAATGCCCTGAAT  
TCCAACCTGACCTTNTACAGAAAAGTGCTTGACTGCCAAGGGGTNTTNCCAATCATTNATGAGGCTC  
TTGGAAAANTCTCCATACTCCTCTTGGGNGANGGCATTAAGGGTTTTNGNCCAAANAANCTTGGG  
CTTGTTTAANTNTTTCNGAACAAAGGAATTTTTCTTTGANCTTCTTANGGGANCTTNTNTTANNC  
TTACCCGGTNGGAATNTTCTTNTTGNAAACCCCAATTTTTNNACTTGGGCTTTTTTATTANTTT

SEQ ID NO: 296 CGAGGTACATCATTGAAATCTTTGGTCTTGTTATTGGAATATTCTTCACGTA  
AGTATATCATAGCTAACTGAATTTATTTCTAAGTATTTTTACAGTTTTATTTTCATATTTTGACATTGT  
GAATTGGTTTTTTTTCTTCTCATTTGTAATTAGCTATTTTTGACAATATAGACAACCTATAGATTTTTG  
TATTTTTATCTTCTGTCTAATGAGCTTATTTAATCATTGAAATTACAGTAGATTTTTAAGAATATAA  
TATTTTGGGTTTTCTAGATAAGAAATTATTTACCTGAACATATAGATCATTTTATTTCTCTTAAT  
GTTTATACTGATTACTGGTACATTAGCCTGGATTTTTCAAAAACATGGTGAAGAGTGATGACAGAC  
CGTGACTGNCTTGGTCTTAATTTTCATGGAAGTAAGAATGCCAAATATTAATAGGGAATAGTATTCC

CTATTANGAATGACCATTTACTTTTGGGTATTAAGNANGGAGCCATTACCATGNTTAAGAAGTTCC  
 CTTTTCTTGTATAGGGGTAATGCTAAAAGGGGCTGGAATTTATAAANGCCTTTCAGCATCTA  
 TTGGTAAATGGATGATTTTTTTCTTTAATTGGTGAAGNAAGATAAANGGGAGGATTTGANGG  
 GAACCAACTGGNTTTTGGACAACTACTTGGGCTTGGGAAAA

SEQ ID NO: 297 GGTACTTTNTTTTTTTTTTTTTTTTTTTTTTTTTCACCAANAAGTATTTATTAAGC  
 AATNTTNTGTGCTAAGTAAAAAAGGCTNTTGCCAGGCGCGGGGGCTCACACCTGTAATCCCAACA  
 CTTTGGGAGGCCGAGGGGGGNGGATCACTTGAGGNCANGAGTTNGGAACCAGTNTGGCCAACAT  
 GGGGAACTCCATTTTACTACTGNTTGAGGAGCCGGTGACANAANCNNTACATCATCTTCATTTTC  
 TTTATTATGCCCAAGAGGGTNCAACAAAAAGGCACTACTATGATTTGGGGGNAACATATTCATCG  
 ACATGTCTATTTGACTGCCATGGATACTGGGGANCAAGGACATTTTGGAAANTTTCTNTTNAANGG  
 ATCTTCTGGGAAATGGGCATTCCCCCNNTTANTGGAANGATTGNTNTTTNACCAAAACCCNTTTT  
 TATTTTNAATAAATTGGGAAGGGNNTCCGGGGNCCCTGGANCCAGGCAAGGGGACCAAAA  
 NAATNCTTGCACTGACCTATGGCAACCTAAATATTTTNAAGGAAAAACACCTTTTTTGNNTTNT  
 ATTGATTGAATTTCCCTTCCCTTGGAAANAACGGNGCNAAATTGGCCGGTTCCTTACTGGTTTGGACC  
 GGGGAAAAANAATGGTTNTTTNTTTTTTTTT

SEQ ID NO: 298 GGTACAACCTTCAAACATTCCAGTTTTTATAAAAAAAGGGGCACACAATCGT  
 GGTTTTGATCCCCCTTTGTTTTTGGACAAATGTTTCTACAAATACAGATTTCAGCAAACCCAAAGGC  
 TGCAATTTACTTGCACAGTAATCTGCCAGCCCATTACTCCACTTAATCCAGCTGAACCTGAAAT  
 GCAAAACCAATGTCCATGGTCATTAGCAATCATAGCAGGTAGAAAGGCTTTATAAGTCCATAAAT  
 GTGCTTTGAAATTCACATCAAATGACTTTTCCATAAGCTCATCTGGACAGTCAAGGAACCTTTTGC  
 CTGTTACGATTCCGGCATTGTTGATTAGGATGGAAACATCGCCGACTTCTTTTTTAACCTGGTCCGGC  
 TACTCTATACACTCCTTCTTTTGGCTGCAATCGCAGGTATAGGCGTGCACTCTTGTGGCTCCAGCT  
 TCCCAGCCATCTTACATGTTTCTCTATTCCCTCTTATGATATCCCAGAGAACAAGAACAGATC  
 CACCCGGCNAACTGCAAGGNTAAAACCTTCCGAGGTCACTTTCACACCTGTGATGANGGCTATTTN  
 ACCACAACGTTNTTCCGGGNTTTNGGAGNA

SEQ ID NO: 299 ACTGTTGAATTTGGTTCGCNAATATTTGTTGAAAATTTTTACACCTACAATCA  
 TCAGTGATATTAGTCTAAAAATTCAATTTCTCATAGCNTCCTTACGTGGCTAGNGGNATGAGNTGT  
 AATGCTGGCCCNAAAAACGTGTTCCGAANAGTTANCNCCCNNTTCANTNCTGTTAAAGACAACG  
 AAGATT

SEQ ID NO: 300 ACAATTGAAAGCAGAGGCATCCTTGAGCTTTAAAGCATTGAACAAACTGGAA  
 AATGCAACATACCACATAACTGAAGTGAAAAAAGTCTGTGTTTTTGTGTTTTTTTAAATAAAAAATT  
 TTCAAAAAGTTAAAAAAGACATATAAGGTTGATTAAAGGGAAAAAAGGCTCCAGTTTGT  
 ACAGGTTTTAAAGTTCTGCTGTGTGTTCAATTGCCTTGTGTAACCACTTGTGTCNCCTTANGGCCAGAT  
 TCCCCTCTCTATCCCCTTTTTTAAATGTCCATTTTGCTTGCCTGGAATTTAAATNTCCTTCCGTCT  
 CANAAATCACANNAAACTTTTGGTTTGGTGACATACANAG

SEQ ID NO: 301 ACTTTTATAGGCAACACCATTCCAGAAATTCAGGATGAATGGGGATATGCCC  
 CATGTCCCCATTACTACTCTTGCGGGGATTGCTAGTCTCACAGACCTCCTGAACCAGCTGCCTCTTC  
 CATCTCCTTTACCTGCTACAACACTACAAAGAGCCTTCTCTTTAATGCACGAATAGCAGAAGAGGTGA  
 ACTGCCTTTTGGCTTGTAGGGATGACAATTTGGTTTTCACAGCTTGTCCATAGCCTCAACCAGGTAT  
 CAACAGATCACATAGAGTTGAAAGATAACCTTGGCAGTGATGACCCAGAAGGTGACATACCAGTC  
 TTGTTGCAGGCCGTCTGCAAGGAGTCCTAATGTTTTTTCAGGGAGAAAAGCATGCAGAACAGATA  
 TGT

SEQ ID NO: 302 ACTTTTTNTTTTTTTTTTTTTTTTTTNAACAAGCACNTNCACTTTATTGAATGAC  
 ACTNTAAACAGGNGTGTGGGTATAAACTGCTGTNTCTANGGGCAGGACCAAGGGGGCAGGGGCA  
 ANAAACCCCAACGTGCATGGCCNNCNTTGCACANTGGATTGCAAAGGTTGCNNGCTATTGGGCAG  
 GCTACTAGTANCCCCGNTTTTCTGTTTATNNTGTAACATANTTTGGTANANTGTCCANAGNCCN  
 GATCCCNAAACAGGATGATTCCCNATGG

SEQ ID NO: 303 ACTTCTTTATACATCTAGTAGACTTNGGCTGTGAATCTGTCTCATCTGATGGA  
 TTTTTTGGTTTGTAAAGTTAAATTTTTTATTACTAATTTAATTTTGGAACTTGATATTGCTTTGTTCTG  
 GGTTTAAATTTCTTCTTTATTCAATCTTGGGAAATTATCTGTTTCCAGGAATTTATCAATTTCTCTA  
 GAATTTTTAGTTTGTGTGCATAGTCATAATAGTCTCATAGGATCTTTTGTATTTCTGT

SEQ ID NO: 304 ACTATATTTATAGGAATAAAAAATTTCCCTTACAGCTTAGTCTTTATTTTCCA  
 AAAAACTTTCTGAAAAATTGAGATGTATCTGATATCCCCGCGCATTTACATGTTGGCAAACATGG  
 CATTTTACTGCCTCTGGTGCACAAAAAGCCCTTGAGCGCTCACCTAATGCTTCTGTGAAGTCAAGG

GTTCCGAGCACGCTGACTCCATGTTGCTAGGGAGCCTCTCTAGGAGGCCAGCAGGCCGCTGGGCG  
 GCCCCGACTGGGATCTGCTGTGGTATGTGGCCCTCTTTGCTAGTGCAGGAGCCTCCTCCGGTCACT  
 TTACCTGCATTCTTTGAGCCTGGGGCCCCAGGCGGGGACCCTAGCAAGAGGCTTACTTACGTGGT  
 CCCCTGATACCAGGCATGGGAATGGTAGCACTTGCTGATTGTGTGGGTCTCTGGCCACAATGGAA  
 ATATTTAACACACACACCTTAAAGTGGAGGTGGGCAGATTTTCTCTGAAAAGGACCAGAAAGT  
 AAATATTTGATTAAATTAATTAATCAATTTATTTANAGACAGAGTCTGGCTTTGTTGCCCAAGCTG  
 GANTGCAGTGGTGCAATCTTGGCTCACTGCAACCTCCACCTNCCAACAATTCTGTGNCTCAACCTC  
 CCAAGTAGTTTG

SEQ ID NO: 305 ACAAAATGGAAGCTGGCCTNAGACAACAAACGCAGCTCCCTCTTCAAGA  
 AGAAGAAGGTCAAGGAAGATGAGGATGATGGTGTGGGTGATGGGGATGAGGACACTGACAGTGC  
 CNTAGGGAGCTTCCGATATTCTTCCCGCAGTAATCCCAGAAACCTGAAACAGACACATGCTCCTC  
 CCTGGCTGTCTGTGATCACTATGCAAGTGGCAGCAGAGTTGGCAAAAAGATGGATAGCAGTATT  
 AATAAGTGGCTCAGTGGCCTCAGGACNGAGGAAAAACCTCCTTTCCAAAGTACTGGTCTGGAAG  
 TTCCANAGGGAA

SEQ ID NO: 306 ACCTACTATGTGTTAGAGACCACTGTAGGGGCTGGGAACAGACATGGAAACA  
 GTGGGCAGAACACTCTGCTCCCAGGAACCTCACACTCTAGAGCAGGGGTGAGCAAAACCTTTTCTGC  
 AAACAGCCAGATAATAAATGCTTTAGGCTCTGTAGGCCACATGGCCTCTGTTGCAATTTACTTGAT  
 TTTGCTGTGAGAGCACAGAAGCAATCAGAGACAAAATGCAAGTTAGCAGGAGTAGCTGTGCTCCA  
 ATATAACTTTATTTACAAAAATAGGTGGTGGGCCAGAGATTATGATCAAGTTTACCTTTGATTAG  
 TCTAATTTTCATGCTGTTGATAAAGACATACCCGAGACTGGGTAATCTACAAAAACAAAGTGGTTA  
 ATGGACTCACAGTTCCACATGGCTGGGGAGGCCTCACAATCACGGCAGAAAGGAGAAAGGCACAT  
 CTTACGTGGTGGCAGGCAAGAGAGGATGAGANTCAAGTGGGAAGGGAAAACCCCTTATCAAACC  
 ATCAGCT

SEQ ID NO: 307 ACCATCTCACTCAACTCTTGCAAGAACTCTAACGAGACTGGTATTATTATTCC  
 TATCTTACAAAAGAGGAAACCGGCCAGGCGCAGTGGCTCACGCCTGTAATCCCAACACTTCGAGA  
 GGCCGAGGTGGTGGATCACCTGAGGTGAGGAGTTGGAGACCAGACTGGCCACATGTTGAAAC  
 CCCATTTCTACTAAACAAAAATATACAAAATTAGCCAGCCATGGTGGTGTGGGCTGTAGTCCCA  
 GCTAGTTGGGAGGCTGAGCCAGGAGGTCACTTGAACCCAGGAGACAGAGATTGCAGTGAGCCA  
 AGATCACGTCACTGTACAGAAGTTTTTAAACCAAACTGAGGCATAAAGCAGAAAGAGCAAAGAC  
 ACATGAATACCCCTTCTTAACAATCTCTTCTACTTATGCCTCCACCGTAACCTTTGTACCCAAACCT  
 CAATCCCCAACGCGGTCTCAAGTTCAAACCTGGGCTCCAGCTTCTGTCCACAGCCACCCCAACATTT  
 TCTTTTTGTATTTTGTCTGCCACTTCAAAAANAACCTTCCA

SEQ ID NO: 308 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGTCTNACTTTAGGTAAGTGGAAAGC  
 AAAGGTGTTCTGTTAAGGGTGACGGTGGGACGGTCCTTCCAAGCTCCGTCCTTGTGGCCTTNTCCT  
 TGTCGCACACACCCCTGAGGCATCTNTGGGGAACACCCAGAACAGACATCCTATAAACATTTTTTN  
 AAGCTGTCTTTATATTTTAGCAATCTTGGGAGGAAGTCNTTTGCTTTACAATAAACCACTAANAA  
 ACAGGTTTTGTGATAGCAATCCTNTTATATAAAAGCAGTANTTNAAACATGGCNTACANGACATT  
 GAGCCTTNGTGGNTATGANCTTTGCCATTACANTTCCACACTTTNANAATTGAAAAANATGTGGT  
 TTCCCATGATTANGGTNACTTGANGAAATCTTAATGGAGCAAACCTTTNGTTTNAAGGTCTTTG  
 TGGCACTAANTTG

SEQ ID NO: 309 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTCAGTCCCTTTACCTTTATTAATTTACGTAAT  
 CTGCANAAAACTCTATGAGATGGAAAAACGAAGGTTCAAGGAAGTTAAACGACCTGTTCAAATT  
 CACACAGCTAAGGAGTGGCACCACCAGCATCCGAACCCAGGCTGTCCCACTTGCAAACAACAACC  
 GGAACAATGATGAGAAAACCGACAAGAAAAATAATGACAATGACGAANACCTTTGTGCGCANACC  
 ATGCTCTTTTACTCTACACAAAAACCTTCTGGGANAGCTATGTCATCACTCCTACTTTACAGATGA  
 GCACACTGAGACTGAGAGAAGGTTAGGGACAACAGCCAAATGCCCCAGGAGATGGAATTCAGAT  
 GGAGATAATGAACAGCCAAGTCCTCCGTAAAGTGAAGTATGTCATACACAGTAAGGTGTTATTAT  
 TAAATCCCTAGAATTCATGGAAAGATGTTCTCCCTCAAGTGGTTTCATGTTAATTCACCTACAAAC  
 ATTCCAGATGCCCGGTACAGGTTTCACTATTCAAATATATGATGTTAACTAACAACTCATGAC  
 CTTCAAAGATGTCTTCGTCCCACGCACACACATTTGGAAATTTGTGTCCATTGCTATTTCCCTTCTT  
 CTATAAT

SEQ ID NO: 310 ACCTGCAGGCCTCCTACACCTACCTCTCTCTGGGCTTCTATTTGACCGCGAT  
 GATGTGGCTCTGGAAGGCGTGAGCCACTTCTTCCGCGAATTGGCCNANGAGAATCNCTAGGGCTA  
 CNAGCGTTTNCAGAAGATGCNNAACCAAGANTGGCTTCCATACTCTTTTTNCNAAACATTAATAA

SEQ ID NO: 311 ACAATCACATTTAAAAATAGAGAGAACAAGCCAGGAACAAAACCCCACTAA  
CTCAGTATATCTAGTTCAGATGTGGCTAATGTCCCTATTTAGCATCTTCAAGTTGCTCACATTTTCC  
TCAGCAAACCTAAGAATGCGTCTCTAGCACTGCGGAGCCTGCCAAGTCCAAGGACATGGGAATGTG  
AATTATCAAACAAGCTGGGGATGAAAGGTGCTGAAATCATCTTTTTAGAAATATTGCCAAAAAGCA  
GCTTGAGTTTTTAGGGTGCTAATTGCTGAAATATGTTTTAACTTCCGAATTCATTCTAAGCTTGTGA  
CCCACTAACCACAAAAGTAAGAAGGATGTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAATGTACGCG  
GGGGCTCTCGCTCGGTCTTTCTGCCGCCATCTTGGTTCCGCGTTCCCTGCACAAAATGCCCAGCGA  
AGCCCAGAAACCGTCCCTGCTACANACAGGAGTTGCCGCANCCCCAGGCTGANACAGGGTCTGGA  
ACAGAATCTGACAGTGATGAATCAGT

SEQ ID NO: 312 ACGCGGGGAGGCTTGAGGGAAGCATGGAGGTCCATGGCAAGCCCCAGGCTA  
GCCCCGANTTGTTCGTNCCCCACCGGGATTCTCANGANTCCNAGTGTCCAAGGAGCTGCTTGACG  
GNGGGAAACGCCGCCCCNGATGTTNTGGNACAGGTTTNTCATCANCTCCNAACCTAAGTNCAGAA  
ANACTTCANTCTTCAA

SEQ ID NO: 313 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGCGGAGGCAGGATCGCCTTATATTGCT  
CAGGCTGGTTTTTCTTCTTTAANAAGCATTTGAAATTGTCATATTTATGTTGAAAATTGTGTATTTG  
ACAAGCAAATACACAAGCACAGTTTCAATTAAGGGGTAGAGTCCATTCTTTAGTTTTAGTTTAAA  
AAAAGTGTTTAATACACTTTTAATACTTTTATATAGACCAATTAAAGAGATCAAGAAAGTAATACA  
TGATATAATAATTTTGTAAATTGCTAGGCCTACATGGGAGTGTGTTAGTCCATTTTGTGTTGCCATA  
AAGGAATACCTGAGCTGGGTAATTTATAAAGAGGTTTATTTTGCTCACGGTTCTGCAGGCTATACA  
AGAACATGNTCCAACATCTGCTCCTGCTAGNCTTAAGGAAGCTTCNATCTNGCTGAAGGTGAAGG  
GGGCCAAAT

SEQ ID NO: 314 ACACACTATATTTACATCACCCACCCTGAAAACAGCAGGTTCTGGCTTTTCCG  
TGAACCCCCAGATGAATATAAATTGGGAGCCTTGAAACAGTTTCTTTCCCAAACCGGGAAGCGG  
TTGCTTTTGGGCCCTTCCGCTTTCGTATNNGCTCGTCCCCTTGAC

SEQ ID NO: 315 ACGCGGGGAGGGCTTACGTGGTCTCATGTTGCTCCCATTTTTCCTACTTGT  
CAAACAACAGTGACAGAACACCTGGTTCCCATGGGCAGGGAGTTCTGCTGGGCACTAGGAACGCT  
GAGATCAAAGGGACCCTGCGTTGCTCTGAAGGGGCTGACAGTCANAGTAGGGGAGGCANACACC  
TAAACNGGTGTTACNCTTNATANCACTGAGCACACAGGGCTCTTAGAGAATGCTNCANGCTNGN  
ACTAANAACCTGATGCCTGATGANCAANGGACNTTTGGGATGC

SEQ ID NO: 316 ACGCGGGAAACCAGGAAGATACAGAATCTCTGAACAGACCAATAACAAGCA  
ACAAGACTGAAATGGTAATAAAAAAGAAAATGCCAACAAAAAAAGTCCAGGACCAGATGGATTC  
ACAGCTGAATTCTATTAGACATTCAAAGAATTGGCACCAATCCTAATGACACTATTCCAAAAGATA  
GAGAAAGAGGGAAATCCTCCTAAATCCTTCTGTGAAGTCACTATCACCCCTAATCCAAAACCAGAGA  
AGGACTTACCACCACAGCAAAAAAAAAAAAAAAAAAAAAAAAAAGTTTCGCGGGGATGTCATCCATT  
CCGANACTGATGCCTGGGCTTCAAGGACCNTTAGGAAACATANACTTCTCCCAAATGGAGGTTTG  
GATCACCTGCTGCCCAAAACATANTGGGAGATTTCAAGTNTTTAATGTNCCACCAGANGTCTNNTC  
CTTGCTCTGCCCGCGGCGCGTTCNAAGGGC

SEQ ID NO: 317 ACTAAAGAGAAAATGGGACTACAGATCCACCAGAATGACTAAGCCAGGTGC  
AGTGGCATGTGCCTGTAGTCTGCTGCTCATTGAGGAGGCTGAGGCAGGAGGATCACCCGAGCCCTGG  
AGTTCAAGGCTGTAGTATACTAATGGCTGTGTCTGTGAATAGTCATTGCACTTCAGCCTGGACAAC  
ACAGTGAGACCCTGTCTCTCAAAAAAAGAAAGACTGTCCCCAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 318 ACGCGGGGGGGTTATACCTGTCTTGACGCCATCCGAGATCACGCTTCTGTCTT  
GTCTTTCTCGAGAAGCAGGTGGAACATGAGCATTGAGTTTCTTGGTACGCGGTATTATTATTAT  
TAGACGGAGTCTCGCTCTGTCGCCAGGCTGGAGTGCAGTGGCGCTATCTCGGCTCACTTCACTGC  
CACCTCCGCTCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGCATTACAGAT  
GCTCCCCACGCCCAGCTAATTTTTGTATTCTTAGTANAGACAAGGTTCTCCATGTTGGCCAGGC  
TGGTTTCCGACTCCTGACCTCAAGTGATCCACCCGCTTGCCTCCAAAGTGCTGGGATTAACAAGG  
CGTGAGCCACCCTGCAGGCCGATGCTTTCAGCTTATTAAGAGCTCACCTTGTGGATGACTGCANA  
ACATTTTCTTGGTAATGTAAACCAAAAGCTTCTGTATTTCCCAAATANCCCCATTAATCAACATTTA  
CAAATAATTGTAAAAATAATTTCAAGAAAATGAGAAAAGNTGACCCCCAGACTGGGGAAAAACTAT  
TTACAAAAAACTTATTTAATAAATGACTGGTTTTCAAAATATTTAAAGAACCTCTTTAAACCCAAG  
AATTNGAAAANTTNCCCNGATTAAATNATGGGGCAAAANATCNGAACAACTCCCCAAA



SEQ ID NO: 319 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTANANAACACACATACAACAA  
TCATGTGTCTNTTAAAGCATGCATTGCAANACAGCATATTTGGTATGTTGGATATCTTGNGTTTTT  
TCTCTAACTTCACTTACCACTCTCACCACCCTGCTNTGGAAATTAGGACTNTGATCTACCCAAGCA  
GGCTTCCTNTGACTTCTGGTTGAGTTGGTCAATGGGGANCACCACAGGAGATCAAAAGGAGGAGG  
GAGAGAGAANACAGATTATTTATTGCCCANGCCCTTTTCTTGCAAGGTGGGGGANAACATTTTGGGT  
GTGCTGCAACCCTTGACCCAAAGTCACAAATACTGTCAAAAAAANATAAAAAACAAGTAGTGGGA  
AAAAAATGGATTCAACTCTATGTCATTTAAATGCTGCTGCATTTTAAATGGAAATGT

SEQ ID NO: 320 ACTCAACAGGCTCCCTAATAGATACAGAATCTGTATGTTCAATTGCATTTCAA  
TTCAAGTCTCCTTATAACTTCACACATTGTCTAGCGCATACTCCTCAAACCTCTTCTGTAGTTTTTTT  
CCTTGAGCTGCGACAAGCCAGAACATGTGTTGCTTCATTCCCTAATTGCTGTTTACAGATTTTGCATG  
CTGGTATTTTTAAAAATTACACTGTCATTAACCCGAGGCAGAGGTCCAGTTGGAGCAACAGGGTG  
AGGAATCTTGTGCAGTTGTCTCAATGGATTTCTTTGCTCCACAGTTGAATGCCAATTCCAGATGTGT  
TCTTTATCCACACCCTCGTCTTGCAATCTTCGTTTTCCCAAATCCATTTTATTTCTTACTCAATT  
AAAAGAATCCAAATCCTTTNTTACCCAAAAAATAAAAAAAAAAAAAAAGGA

SEQ ID NO: 321 ACTTTTTTTTTTTTTTTTTTTTTTTGGGTTTTTTGCCAACATACCATTTTAAATGGANA  
CTCAAAACATTAAAAAATAAATCANAACGTGANCATTGCCAGNAAAGGTCAAACCTGCCATAGG  
ATAAACTTTCTGGGTCTCATATGAAGCCTTTACANACANAAGCGTGTCTATGTTTCATGGCCTTTA  
TGGATGNAACTGGA

SEQ ID NO: 322 ACTTTTTTTTTTTTTTTTTTTTTTTGNTTAGGATGGANGGGAAATTNTGCCTGTN  
TTCTGTGATCTAAAAAACTGGTGCTGTCCANCCTGCTGAAAAATATTTNANTATNANCTATTAC  
AATATATTATTAGGCTTACAGGCTTAGTGNTAATCTGTGTTAAATTTCTGAAAATACAATNATTAT  
TCCTTTGCTNTAAGGAAA

SEQ ID NO: 323 ACGCGGGGGGTTATACCTGTCTTGCAAGCCATCCGAGATCACGCTTCTGTCTTG  
TCTTTCTCGAGAAGCAGGTGGAACATGAGCATTCAAGTTTCTTGGTACTTACCATTTTGAAACAT  
CAGCAACTGAAAAGTGGGTACATGATACAATGAGTGATGTTGCCTTTCTCCAGACCTGTTCTGTT  
TCCTTGCCCTCTCTGGCCTCTTTCTGGCTCCATTTACCTTTTCCACTAGTTTACCAGATGATTATCA  
ACTTTCTGGCATTTTCCACTTACCATCTGTGTCCAGCATCCTGGTAAATCCAATAAACTTTCTAAT  
ATCCAGCATCCCATTAACATATCCTTGATATAGCAAATCACTGGGAGTGGATAGTCATATGAT  
GGCCTCAAGGCAAGGGCTCTACCACTTTTTNGAATAACCCACTTGTGTGCCAAAGTTCTGTGCTGG  
TTTGCATCAACTAAANTCATCTCACTTTAAAGCTAAANTATTAAGATCCTAAAAAAGCCATTTT  
ACAGAAAGATTANAATATTTTCAATTATTAATTCAGAATAAATATATCTTTTTTCTTTAAACTNCTC  
NAATAACTAATTGAATTGCATTGGTTTNAATTAAATGCNGTCATGTGTATATATAGAATTAATAATC

SEQ ID NO: 324 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTANANACAGGGTCTTGCTCTGTTGC  
CCAGGCTGGAGTGCAGTGACACAATCATAGCTCACTGCAGCCTCAACTTCTGGGCTCAAGCGAT  
CCTCTTGCTCAGCCTCCCCAGTAGCTGGGATTACAGGTGCTCACCACCACCTGGCTCAATTTTA  
AATTTTTTTCGCAANACAGCTATGTTGGGAGGGTGTGTCGACCTCTGGCCTCAAGCGATCCTC  
CCACCCACCTTGGCCTCCCAGAGCACTGGGATTACAGGCATGAGCCACCAGGCTGGTCTGGACT  
GTGTTTCTATAAAAAGCCGGAAGCACACCCAGGGGTGCTCCCTTTGATGACTTTGCATNTTTCTCC  
CCAGATCCACCCACCTACCCCATAGGGCTTGTTAGATACAAGGACTTAACACAAACACAAGCTG  
GGGACAGCCTGTGAGTCTTGATTCTGCAAGGCACCCAAGGAAAGTATGGGAGCTTTCTGCTGTCTC  
CTCCAATCATTCCTNTANCCCTTTCATGGCCTGGTAAACACTTT

SEQ ID NO: 325 ACTCGGGAGGCTGAGGCACGAGAATCACTTGAACCCAGGAGGCGGAGGTTG  
CAGTGAACCCGAGATTGTGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCTCAGAAA  
AAAAAAGAAAGCAATCAGAACCAGGTGAGGAGGAGAAAGGCAAGAGTTGCCACTGCTCTC  
CTTCTGGGATCTTCTACCCAGCTCCAGAAAGGCAAGGGGCCTTGGATGCTTTGGAGCCACTGNTC  
AGCCCCGGAACCATCACCTGCTCCTCGGGTTTTTTCATANAAAGAGACACATNCTGGAAGAG  
TGTTG

SEQ ID NO: 326 ACTTTTCTTTTTTTTTTTTTTCCAAATTTTACCTTCTTTAAAAATAGGTTGTTGGAG  
CTTTCCTCAAAGGGTATGGTCATCTGTTGTTAAATTATGTTCTTAACTGTAACCAGTTTTTTTTTATT  
TATCTCTTTAATCTTTTTTATTATTAAGCAAGTTTCTTTGTATTCCTCACCCTAGATTTGTATAA  
ATGCCTTTTTGTCCATCCCTTTTTTCTTTGTTGTTTTGTTGAAAACAAACTGGAAACTTGTTTCTTT  
TTTGTATAAATGANAGATTGCAAATGTAGTGATCACTGAGTCATTTGCAGTGTTTTCTGCCACA  
NACCTTTGGGCTGCCTTATATTTGNGTGTGTGTGTGGGTGTGTGT



- SEQ ID NO: 327 ACCTAGGGAGTGGCAGAGTAGTGATGTAAACTCAGGTCTCTATTACTCCTCG  
GTTCCCANACATCTTCCTCTTTGTCCCTCTGTTTTTCATAAAGAAAGTCTATTCCCTCTCCACNANCTC  
ATTTCTTTTCCTGGGACACAACTGGGAAACNCCTTCTTTACTCTNGTNTCTCTCCCTTTTTATAC  
AATTATTCAAGAGACTCATCNAACTTNGCNGGGGTGGCTGTGACATANCTCCNNAGTGAAAAC  
TTGGAAGAGACATNTATNCTCATNCTGTATGNAAATGTTCTGTTAAAATGTGCAATCCTGNATTA  
ATCCTGGGCTTTCTAAGGCGGNANAATACCTTCATAGCCTCAGTAGCTGGCATGTGGTAGGTGCTN  
CGTTCNGTTTTTTNGTTTGTGTTTTGGTTTTCTGTTTTTCATTGGCTAGTTTTATACAGGGATTGT  
AAAACCGCATAGGCCCAAATCATGCCAAATGT
- SEQ ID NO: 328 ACGCGGGCTACTCCTCTGAANCAAGANGGAATTAACAAAAGACAGGCCAAAG  
AGACAAAATATTGANCAGGTGTTGCAANAAAAAGTTAANGGCATCNTAANNNGCNTCTTTNANA  
AAGNAAAATTCNNTTCCATTTATTAATNTTTATTGAGACTCACTGTGNTNCCATNCAATGTCTACC  
CCTGAANNAATTTCAATTTTGGCNAAGATTGTANTTTCTGTTCNTAAGGCNTATATTNTTTTTANA  
AAAAATAATATCATANTGAATAAAGGGNNGAAAAATAAATTTTAACCACTNTNNGGGGATACTTT  
CAAAAGGGGCTCTACNAAAAATAGGCATTTACATNCCATNTTGAAAACTCAAAAAAGTTTTCTT  
AGGATCANCTNGTATTTGNGAAGAAAAACCCNCTTACCCCAAAAACTGACATGGGTNGCNTGGA  
CCTAATAATAATTTTTTTAAAAANAACCCT
- SEQ ID NO: 329 ACNAAATTTAAATTAAGCNACTTTCTTCCGNATTAANAAGATTCTTCACC  
TNCCTAAATTAGAAACCNAAAGGGGTAAACNNTNCTATTTAAGGATCATACNNTAANGCTTATAAA  
NAGGCCNCCAAATTAATCTNTTAACCATTTTTCTTAACNCTTTNCTAGGGGACCTTAACATTTTAAT  
CCCAACCTNNACTTGTTTGGCTTGGTNTTCTAAANCTAACAATATTTTTTAAAA
- SEQ ID NO: 330 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNNTNGGCAAGGGCAATTCACATTTATTT  
CCTGAGCATGCTGACTAAACATACTTCAAACACAGCAGAAGGTAAANAGTCTGTTTTGTAAGTAA  
AGCCGTTTCTTGCCAACAGCGAATGGTGGTCCATGAAAAGTACTTTTTTTTTTTTTTTTTTTTTTNA  
AACAAGGACAATNGTTATTTANTTTTTATTTCAANAATCANAAAACTNAACTTGCAATCCACNTNG  
GCNTGGGAGGNAACAAGGAAAACNTGGGAACCCAAAGGGAACCTTG
- SEQ ID NO: 331 ACTTAAATGAAGCATATTCATGTAATGTGCTTTTTTTTTTTTTTTTTTGGCCAGC  
TTTTCTAAGCAAATAGATTGTCTGAATTAGTCACAGAATAATTTGTGAAAATTCATGTTTAAAGTA  
GCAACTACCCTTTCTTTTTTNATATTTTTAAGGTATTAGTTTATCTTCTTAACCTGGTGCAAGNCA  
CTTAATGTTTTCATTAATCT
- SEQ ID NO: 332 ACAAGATATAATTGATAAACCTGAAAATTTAAAGACACTCCGAGTGAAGAAA  
ACTGAAGTTTATTGGTCAATGGAGACAAACACAAAATGCTATTACAAATTCAGAAAGGTCCCAAG  
AGTCACTAAAGATTATTTTTGTGAAAACAGGTTATACAGCAATGAGAAAACAATTTAATACATTNT  
TTCCCTCAAATGAGATCCCCTCCAACTCCCTGCATCACTATTTTTCTAATTATCAAACTAAANAC  
TTGTGAGAGAAGATTGAAAATANAACCTTCAGTTNAAAGCTCTTTAAAGAANCACA
- SEQ ID NO: 333 GTACAACACTAGTTGGAAAATGACTTGGTAAAGCAATTAATGTNACTTTCA  
CTAATAAAAGAAAGAATTCTTATAAATAAAACATNGCAAANAAATAAAGGAATTAATACAGCTGT  
CATCTGGCTATCTTACAGCTTTTTTCCAGTCNTATATTTTACACCCACTTTATTANGAAAGCTTCTT  
T
- SEQ ID NO: 334 ACATATTACTGAGACCTTATCTAACATGTAATTGTATCTAATTACCAGACACA  
ACCAACCCCACTGTAGAACTTGGGCAAAACAATTAAGCACACAACCTTCTTCTATATATTAATA  
CTTTTACGTTCTTTTTTATAAAGAAAAAACCTGGTTAGAAATTAGGACATCTAGAGCCAGACTT  
AAAATTTTATGAAACCAATGGGCCTGACTGATTTAATGTCATCTCATAAATTATCANGAATTTTTA  
TGTTCTCACAAATTAGATTGTATATTTGTAGTTAAACTCGGTCTACCTTCAATTTCTGATCATT  
TGCTAGGAAATAATAACTGGGAATTCAACAATATTTTAAAAAGGTCAAATAATGGTCTTATCAA  
AAAATAACTGAGCCCCAAATTCAAATCACTGGCTAACCTACCTACCCCAATGGGTAGATACTTN  
CAAAATTTGGNATATTTTAATTINCAAGATCATTTTAGATGTTT
- SEQ ID NO: 335 ACCGATCCTGAGACCTTCGTGCAGGCAATCTCTGATGCCCGCTGTGTTTTGA  
CATGGGGGCTGAGGTGTTTCAGCATGTATNTGCTTGATATTGGCGGTGGCTTTCNTGGATCATN  
ANGATGTTA
- SEQ ID NO: 336 ACAAATGAGACAAAGGCACAGAGGTTAGTTCACATAGCTATGAGGCACAGG  
CAGAATTCAAACACAGGCAGTTTGGCTTCAGAGACCATGATCTTAACTGCTATGCTCTGATGTCTC  
TCCAAAAAGTATAAACATGAGCAGGGTTAATTGTAGCAGCTACTTGGTTTTTACGTCAAGAATCA

TAAACCACAAGAGGAAACATGAAGTTTTTGTCTTTTACTTTTCAAGATGGAGTCTCGCTCTGTAC  
TCAGGCTGGAATGCAGTGGCCCTATCTCAGCTCACTGCAACCTCACCTCCAGGGTTCAAGTGATTT  
TCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGTGTACACCTGGCTAATTTTTGTATTT  
TTAAGTANAAACAGGGTTTCACCATATTGGTCAGACTGGTCTTGACTCCTGACCTCTGGTGATCCA  
CTC

SEQ ID NO: 337 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGNCAATCAACAAGTGTTTATTGATCACC  
TACTGTGTGCCTGGCACTGTTACANATAGTCTGGGGGATACAGANAGGTCTAGGATATGGCCCCC  
ACCCACCGAAGGGTTTACAATATACTTGTGAGATCGGACACACACACAAATAACGATCAATCA  
AAAATTGTGAATGCTAAGCATCAAAAAGCAATTTATACATTGAGGGTTGGGGGAGGGAGGGGT

SEQ ID NO: 338 ACCTAGGGAGTGGCAGAGTAGTGATGTAAACTCAGGTCTCTATTACTCCTCG  
GTTCCCAGACATCTTCTCTTTGTTCCCTCTGTTTTTCATAGAGAAAAGTCTATTCTCTCCACGAGCT  
CATTTTCTTTCTGGGACAGAGTCTGGGAAACCCTTTCTTCACTCTGTTTCTCTCCCTTTTTATACA  
ATTATTCAGGACTCAGCTCAGACTTTGCTGGGGTTGCTGTGTCTAGTTACTCTCANTGTAAACTTGG  
GAAAGTCATTATACTCATGCTGTATGGAAATGTCTGTAAATGTCCAATCCTGCATTAATCCTGG  
GCTTTCTAAGGCGGGAAAATACCTTCATAGCCTCAGTANCTGGCATGTGGTAGGTGCTTCGNTCTT  
GGTTTTTGTNTTGGTTGGTTTGGTTTCTGGTTTTCAATTGGCTAGTTTTATACAGGGATTGNAAAACC  
GCATAGGCCCAGATCATGCCAAATGT

SEQ ID NO: 339 ACATTTGGCATGATCTGGGCCTATGCGGTCTTACAATCCCTGTATAAAACTAG  
ACAATGAAAAACAGAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC  
CCAGCTACTGAGGCTATGAAGGTATTCTCCGGCCTTAGAAAGCCCAGGATTAATGCAGGATTGCG  
ATATTTAAACAGAACATTTCCATACAGCATGAGTATAAATGACTTTCCCAAGTTTACACTGAGAGT  
AACTGACACAGCAACCCCAGCAAAGTCTGAGCTGAGTCCTGAATAATTGTATAAAAAGGGGAGAG  
AAACAGAGTGAAGAAAGGGTTTCCCAGACTCTGTCCAGGAAAGAAAATGAGCTCGTGAGAGG  
AATAGACTTTCTCTATGAAAACAGAGGGAACAAAGAGGGAAGATGTCTGGGAACCGAGGAGTAAT  
AGAGACCTGAGTTTACATCACTACTCTGCCACTCCCTAGGT

SEQ ID NO: 340 ACACAGAAAGGGAGGTGTCAACAAAAGAAGATAAGCCCATACAGTGCACAC  
CTCAGAAAGCCAAGCCAATGCGGGCAGCTGCTGACCTGGGGAGGGAGAAGATCCTCAGGCCACC  
AGTAGAAAAATGGAAGAGACAGGATGACAAAGACTTAAGAGAAAAACGTTGTTTTATTTGTGGA  
AGAGAAGGGCACATTAATAAAGGAATGCCACAGTTTAAAGACTCTTCAGGAATGTCTAAATCAGA  
TTGTATGTTTGGATCGCCCTCACCTGTCCATTGAAACCAACTGGTTTATTTGTTTCAGGCAGTGCTT  
TTACATGAAGACAAAAAGAAACAAAAACAACAATATTTTGTAGTCCCCAGTCAGGTAGCCTTTC  
CAGTAAATATATGACTCAGGGAAGGCCTCAGCGAAGAGGACCCAGCAGGAATCATGAGGGAAG  
GAAAATGCAGCACTCTAAATGGCCACTCAGGCGTTTCTATTCACTCGGAAAATTAGGTTTCAATTTCA  
CAGGACACAGCAGTGTAGATCAGGCTTCAACTTAACATTTAAGGGAAATGTCAGATTTTTTTTTTAA  
TTAATGAAATTGTTAATGAGGAAAAATTTTTAATATAGTCTTATCTACCAACATCCCCATAGATTT  
AAGGATTTTAATA

SEQ ID NO: 341 TCATACAGACCATGGAATACTNTGCAGCCATGAAAAGGAACANGATCACGTT  
CTTTGCANAGAGATGGATGGAGCTGGAGGCCATTATCCTTAGCAAATAATGCANGAAAAGAAAA  
CCAAATTCCACATGTNCTCACTTATAAGTGGGAGCTAAATGATGANAACACNTTGACACATGTTGC

SEQ ID NO: 342 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTAGANACAGGGTCTCACTATGCTGCC  
CAGTCTGGTCTTGAGCCTCTGGACTCAAGCAAACCTCCTGCTTTGGCTTTCCAAAGTGCTGGAATC  
ATAGACATGAGCCCCCATCCTTGCCCAATTTTAAATATCATATATAAAAATTTCGGACTTTTGTTTT  
AGCTTTCTGTTTTTAAATTTCTCTCAAATTAAGGAGCAGTTGAGTTTGATGACTAAGAGCTTGGG  
GTTGGGACATGGAAGCCAACCTCTCCATCCTCAAATTGTGTCTCATATAGTTTATGAAACTACCCAC  
AGACTCAACATNATAATATGGAATAATGTAATAACAAATACATACTATCTATAAAATNATTAGCCA  
TATAAAGAAAAACTTCCACTCCCTCTCCTTCTTCTTCTGCTCAA

SEQ ID NO: 343 ACCTGGNGGGTCTGTTTCAATNTNCCAACCTGTGTGCCATCCACGCTAAGAG  
AGTACCATCATGCCCAAAGACATCCANTTGGCTCGCCGATACGGGGAGAGAGAGCTTAAGTGA  
AGGCAGTTTTATGGCGTTTTGTA

SEQ ID NO: 344 ACTTTTTTTTTTTTTTTCNNGGNTTTTTTAAAGTANNGGNGTNGAGCCCGAACG  
CTTTCTTAATTGGNGGCTGCNTTTAGGCCTACTATGGGNGTTAAANGNTTTACTCTATCTACAAGG  
NTNTNTTCTANAGCCCAAAGAGCTNTTNTCTTAGGACTANCATTTTAAATNTACAANGGGATCCT  
AAAGGGNTATGTGGNNCNAATNTATAGTTTANCGTANTTATTCTATNTTNGNTNAACCANCNTNT  
NANCAAAACACTGGTATNGTTATGNAGCCATCTACCTTTAAATNTTCTCTCNTNTTNTNCTACA

TATTACAGATTGNGCTTINTNTANGNCTGGGCCCTTTTNAANAAANNTTNCANGNGTTTAAATNGNG  
TGATTAACCTCTTNGNANCTNCTTNTACANANCTANTNNCNANTTTAATTNTTTAATNNCTTAAAGT  
AAANTGANGNTNATNNNTTTANTNTTATNAATNNATGGGNAAAAAATNCTCTCAAAACTATNCTA  
NTTTTNNAAANCTTTTTTAATTNCCACNCTAATANGGNNCTNAATTCCATNATAACANTGTNGG

SEQ ID NO: 345 CGAGGNCNCGGGGGGGACCTGGGCTGCAGTCTTTCTATTGTCAATGGCCTAT  
ATGACCTATGAGGTTAGTTAATCNTTTAGAATCTTAGTTTCTTCACTTGTAATAATTAGGTAATTGAA  
ATATGTGATTTCTAAAGCTCTTTTAACTCTAAACGTGGTTCGTGCATANGAATAATTCTGTGCAT  
AGGAAATACACATTAGAATATGAATGTCTTCTCCACATTATTCTGTAAATCCTATTTACNTCCCAT  
TATCCANAGTAAATATGGGATTNCAATGAAGTNTATTAGTTACCTAACACTTTAAGCCTAATGGT  
GGCTTGATTTTAGAAACCACTTTGAATTTTCATGCANNTATGGTATANTCTCCTGAGTCNTTTCCGT  
TCTCGCTNGATTAGCNCCTNTGTGCACNGATTNTNCTGTATTGGCCATTNNATGCCTNNNTGCT  
GNNTNTTAGCCTGTNTGTGCNTCTTCTTTCNANNAACCGTTCCNNATCTNTAACNTCTTTN  
ATACGTGNNTTANTTNTNTGNATTTTNCNNGTTGNNNNTTNTANAATNTGATTTTCNATNTATCG  
GGNGNTCNNCTTNTNTAGGNTGANTCTTNTTTGTNNATTATTANNTNNNTNTCNCG

SEQ ID NO: 346      ACTTTTTTTTTTTTTTTTNTTTTTTTTTTTTTTTTACGNGGNTTTTTTTTTTTTTTTT  
TTTTTTTTTTTTCTTNACGGNTTCAANGGACNCTTNNATNGTTTACTAAATGGANCATCAATTTNGNC  
TNACTACCTANAAATGGAATTNCATCTGGGTCCATGCTGANTTGANAAANANGNACNAANNTAN  
TCNTAATAACNTANATTNAAANNANAACCTAANCTAACNCNGCTNATTTCTTTTAAACNGGCTAA  
ATTTAANTNTATCCNTTCATAAAATGCNNAATTGGTTTANTNNNAAAATAATTNTAAANGGNATCT  
TTGAANAATNTNTTGCAANTNCNNGGAGCTTTGAANATTNANNTTTAANANCTTGNCTTACTGTCTN  
TAAAGGGTNTTNANNCATATGNNTACNNTCTGGCCANNNTCTAAATTTTNAATTTCCCTATTNAACT  
CTNTNTTNTTATCTTNCNTTATGTNTNAAAGATNGACNNANTNTNNGNTATNACNTNACTATGTNG  
NNNTNNTANCTACTTTCNACNTATANTTCNNTNAAANTTCNGTACNTTTATNNTGTTCNTTTATAAA  
ATNNATTTTCTTAATATGNTTNTATATCCANTATNTNTCCTTTNNNTTATTATNTGCNTCNG

SEQ ID NO: 347      ACCGGGGGTAGTGNCTTATTGCAGATAATTTTINAGCNTAGGGNCTGGGGGNT  
ANGACGNNTCTCTCNTTTTINAGTCGGAGACCTCTGCNGNATACTNGCTCCNNCANANCATNTGCN  
AGGCCATGAAGCTTCCCAACTTNTTCCNCCTNTTTTANTATTATTGAACTGGNGCNANTCGNAGCA  
ANNNTGCATNTCTTGCNTTGNCATNTGATTACTCCAGATNTATTAC

SEQ ID NO: 348      AATNCGCCCTTAGCNCGGNCCNGGCCGACGNACANCGGTACCGCANCATGG  
GCCANAATGTTGCATATTACATGCTCTACTNAGTGGAAGAAGATGAANATGCCNACAAGAAACAG  
NTCGCTNAAGTCTNTGTCTGAACANNANGAGCCCCCGAACGTGACTGN

SEQ ID NO: 349      ACTNTTTTTTTTTTTTATTAGANTTNTTNCAGACACACTTGCTNNGTNTCCAA  
GGTTGGAATACANGGGCATGATCTCAGGTCAGTCAACNTCTNCNNCCAAGGTNCAAGNAATTNC  
CTGGCCTTANACTCCNAAAAAGNNGGGATTACNGGCNTGTTCCNCCNCACCNAGGTATATTCTNG  
TATTNTATAGAAATGAGGNITTTACCATATTGGCCCCGGCTGGCCTNAAATTTCCTAGNCTNNANCAA  
TATANNCNACCGTTGNCNNCACCAAAGNGANGGAGNNTACNGGCATTTAACCAANTGCCNCNGNCC  
TGAANANCNTTTTTNTGATTTCAANNCTTTNTNTNGGCCCTGGCGANTTCTCCTTACCCGNATNNAT  
NTTTTNCNCNAAAAAGAAATAAAAAACC

SEQ ID NO: 350      TCGAGNGGCCGNCCNGGCANGTACATTCAAAAANCNTNAGGAAATATTNTG  
ANTGCCCANGNTGATGAAAACCTGGGGTGAATTAACCTCCACACATTTTATTTCAGNNTGNTAAAG  
AGNTTTAGNGGNGCCAACGCAATGGTCNCTATGCATGCNNAGAAGNCAAAAGAGNTNCTCC

SEQ ID NO: 351      ACTGCATAGATTAAAGAAATCNACTGCNGNANNCCNCTCGTANGGAANGA  
ACGCCATTGCCAATGATAAGNCNTTGCACNTAGGNNNTGANNGCAAACAANTATAGTGNGCTTCC  
NAACAGGTNATAACCANCTGATAAACACCATTANANNNCGATGCCAAGCATGTNNCNCATNTGG  
TGTGACCAANNACTATTCATANTGAACAAAAGTTGTGTATANATNCATNCNAAGGGCAAACCTCCC  
TCCNATGATAATACCAGGCTAAGGGCTTCCTTAGAACNGNTGTTATGGAATNTNACNGNGAGAAN  
CTGTTGGNTCNGAACTGGGNGNTCNAATGGTNTTCNTGGTAAACCNAAATGTGCCCTAACANACT  
CA

SEQ ID NO: 352      ACGGTTCTTCCTGTGTCACTNAATAGCTTGCTGCTTTTTAAGAACCAAGAAG  
CNGNNNAACTTTGCGGCAGCTTGTTTTCTGTTNNTATTACTACATNNCTTATNGNNNACNGTTANA  
TCTANTCCACCTCCGAGCATNGGCTC

SEQ ID NO: 353 ACGCGGGGGCTCAAAGNNGGCGCCATCCGGGACCGGCGGTGTGTCTGTGGCCG  
GAGGCTGATCAGGATTATTGCNTCGAGCTTTTAGNGTCTTNTTATTNANCNCCGAAANTAACCTTN  
TGNTNCANCAAAAATCAAATNCTAAAATTNCCTTTCCAGGTNGTTTNAACAANTNCTNGTTGTATNC  
ATCCNTTAAACCAANCCACCCGCCTTGAGGAAAGTNACCCCTTGNANTGAGGCAANCAAAAAA  
NCGGCTNAAACCTGANCTAGNAATTCCTTNGAAAAAGGTCTCCAAAAAAATTTATTTNTNACN  
CNAATTCCTTAAANNCTAANTCTNNTGTNNTNTGGGGGAACATAAAATTTNTNANTTTTTTGGGG  
AGAAAANTNACCTTTTANTCAAANCCAANGAAATAAACTTTGNTTTTTCAAAGAANACTAAANAA  
CTTTTAAAAAACNCCATTTGAANTTAAATACCTTGGTTAANNNTTGGNGCNNTTTCTTTAAGGGG  
GAAATTAATNT

SEQ ID NO: 354 ACAGAAATTTCAACAAGATGTCAAACACAGTGATGCCATTTGCTATGTTTNAAT  
TTGCTAGTANCTTATTAANCATANCATGCAANTAATCAAAGANAANCNTCCNTGACTTANAGGGA  
AAAATAATTCTANAAAAGTTTCACTAGGTAAAGTNTGCAAATCNTTATNTAAAAACCTTCTTTAAG  
GCCATGAACCTTCATGTNTTTGCAANATTNTTGGCCTCAATTNTCCCCCTTTTAAACCCANCAA  
NTTNGANNGTNTNAANANAATTTGNGGCCAAGAAANTACATCCCNTNGTNTTAAATCCANGTTTT  
GAGGAACNTTANATTTTAGGGTTTATAAACTTNGGCTGATTTCCNGCCACCAGNANAGGGTTTNC  
CATCTTTGCCAGTAAATAAAANNANCTNNTTNTCAATAANCTTTGANTGGNGGCTNNCNGGATC  
TGGGNGGCCNTGGNAAAAAGGCCCTCCCTTN

SEQ ID NO: 355 ACTTGCCTACGGCAGCAACCTGCTGACAAAGAGGATCCACCTCCGAAACCC  
CTCGNGGCGTTCTTNTGTGTGGCCCGCCTGCAGGATTTAANCTTGACTTTGGCANTCCCAAGG  
CAAANCAAGTNAANCTNNGCNTGGAGGGATANCCNCCNTTTTNAANNCCCTGGCAATAAGNG  
NGGGGAGTANTTGGAAAAATGANCAAAANCAANTTTAANTTNTNTGNNTNACCAAAAAGGGTTA  
AANGTGGANTGTNGTAATANAAGTNAAGTTGCAACTCAANAAGGAAAAANAATAACNTGNA  
AANTTTTTNTNACAAATNCAAAGGGCTCCCTCCNNCAGNTAAAAANNTNTTGTGGGNGGCA  
AAAAAANGTTNCCTTGNANTTCAAGAAAGTTAAAGCATTAAACCATNGCTTTNCGGAAGGNTAA  
AAATTGNAACTTTNAAANGGGGAACNCACNTTTTAANATACAAAATTTTAAGGNTNTTTGGC  
TTTAAATTTTACCTTGAACCGGNAT

SEQ ID NO: 356 ACTCTTTGTTTTGGCACACTTTTCTGACAAACAGCCGGTGTCTNAAACACNT  
AANTNCTAGTCCACNTTANCANCANTANTNTGAAACCGCTCTCCGTAAANATNCCAAATNGGA  
TNCAAATGGACTGGAANTNCCTGGNAGGTTTACAAAATTAAGACAAAGGNCAAAGGAACCTT  
GCCAAAGGAAATGGAAAGCANTTNTTAAAAATAGTGGGAGGNAGGACCAANACCTANTAATT  
CCATCCCTTTTAAATTGGANCCCTTTNNCNTNCCCNNGGNTTTNCAANTNAACTTGGTGNC  
CCNACCAGACACCNANTTTAAGTNGTTGCTTTCAAAGNTTAAAGCCNTTAGGGTNTCCTTAAN  
GTCNCCNTCTATGNAAAATNGGGCATGGNCCTTTGACNGCCTCTNCTAAATNGAGGNCANAGAC  
ACTCTATTCTCTCCAATCACNATAAGAGAAAG

SEQ ID NO: 357 ATTCGCCCTTACCGGNGGCCNGNCCGNCGGGCNCCTTGGGTNTGAAGGGGTC  
NCTCTNGCNCTCTGCNANGGGTGCAGCNGNAACNGNTGNGAAGNTTNCAGGCTCTACTATAAATA  
NTGAGCATGGNCNTTAACTGACCAANACTGGNANGANANNGTNNGNTCNTNTGACNNTNCTGG  
GGGCNCCCCAAAANGNANCNTGNGCTTGGGACACACAAATTNAANACCATCACNGGGAATNN  
GCTGCTGTTATTACCCCATCAAGNTGACAACCTGATGCAACAGCANACTCCAGTCTCGCAATAATG  
AAACCAGTGNTTATCTNAGCAGCAAANTTTGTNTNANTCCNTTNANTNTGAACCANACCAACN  
GGAATGCTTANACCATGGGNCNGTCTAAANAAAAATACTTATTTAAATCANCATGTCAACAAAAT  
TAACCTTCNACNTTAAAACTTANGGNNCAACCCCATCTCCAGAC

SEQ ID NO: 358 GCGAGGTACTTNTTCTTCNCTTTCTTTTTTTTTTTGTTTCAANGTNGCNNCNATAC  
TTTNNCCACTNTGNAANNNGTACCAGTTAATNTCTTTGANTGTNGCTAAAAGGACNTTATTGTGT  
AGGGNNAAGTTNTTTTGATATNAGTATAACTGTGAAATNGNAAATGGAAATNTNTCCTTANCNAT  
ACAACACAATGNTGTNCTGCNTAAATCNGNCNTGATTAAATCTGTGCTGTNNAAGGGCTTACCC  
ACTNNNGCNGNTCACTGTGTCTNTATGCTCNATTCAAGNCNTGTNAAGTAGATGTGGGTTATNANN  
GGCANGTTTNNCCNANCATGTNCTNTNNTTGTACATAATANGCCNGNTCNTCANTGNTTGGANAT  
CCCTNTNNAATTNTCNGTCTAGNTCCCAATTNCTTCNCCNATNCNANCNNNCNCACTTNT  
TTTTNGNCGNCTGGNNATAGAACTCTGNCGTGCCAA

SEQ ID NO: 359 ACTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTNTTGAGCTGGANACTNNCTNTGTTG  
CCAGNCTGGAGGGCNAATGGCNTGATCTTGGCTCACTGCAACCTTTGCCTCCCGGGTTNAAACTGAT  
TTTCCGCTNANCCTNCCNAGTNACTGGGATTNTAGGCNCCNCCACCANACCCGGNTAANTTTTT  
TATTTTTTTTTTANTAGANACAGGGTTTACNTTATTGGCCTNCTGGCCTTTGAACATCCTAACCTNA  
GGNGAACCCCTCCTNGGCCTGNCAAAANTNNTANGGTTAAAAGGCGGGANCNNTTGTCTTCT

GNTTGGGATA

SEQ ID NO: 360 ACTTTTTTTTTTTTTTTTTTTTTTTTGGNGCNTTTTTNTTAAAAAGAACNAAATNT  
 TNAGGNCCNAAATNCNTCAATTTTTCAAATGAAAACCCCTTCAAACGGGTNTNTCCTNCATTCAACG  
 AAACCTTNTTNCAAATTACNGAATAATTTAACTTTTTAAAAATAAAAAAATACAANTTNTTAAATGCC  
 TAAAAATTTNTCCCCAAATAAATGNTTCTTAATTTTAANGAANTNTTTTTATGCANNACCTTGGNC  
 GGGACCACCCTNAGGGNN

SEQ ID NO: 361 TCATTTCTACCGAAGACTTNCCNCCGAACNTGTCTGCCAATGAGATAAANTTG  
 CTNTAGGTAAGAAGCTGATGAANNANTGACCCAATANGATGGNCCATTTNTTGGCAAAAGTTCTAA  
 ACAANCNANGAANNACTGGNGCNTTTTCATNTNGACCTCANATTTCTTGCAAAACCCATNACCTT  
 NGANATAAGGCCTTTTTCTTNAATCTTNATCCNNGTNCATACNTTNTCAATGATCTTGCATTN  
 GCCCTTTTGGGCNNCNTAAATATGGGAAATTAANTTTCTTCCANTCCANAGATCTCNCNTNATTG  
 TTCCATTTTANNAAGTTNNCATNATTNGTCNTACCGGTTAAAANAATCA

SEQ ID NO: 362 CCCTTNNCGTAACGGCGCCCGGGCAGGTACTATTANCCATGGNCAACCCAC  
 CGTGTCTTNGACATTGCCGTCCACGGNGANCCCTTGNGCCGCGTTTCTTTGACCTGTTTGCATGA  
 CNNGGTCCCAAANACAGTATAAAATTTGNNNGCTCTGANNACTGNNTAGAAANGATTTTGNTTN  
 TAAGGGTTACTGGTCT

SEQ ID NO: 363 ACTCTGGTTTGTCAATGGGACTTACCAGCNNTCCACCCANNANNTTTTNATC  
 CCAACATNACTGNGAATAATAGTAGGATCCTATANGGAGCCAAACCCNGNANTNATACACTGNC  
 NTNGNNANGACCACANGANCAANAATNTNCCNGNCTATGCTGAGCGGCCCAAGCN

SEQ ID NO: 364 TNCGNATGCTACTTGNNCANTGATGGTAAAAGGGTAGCTTNTCTGGTTGATCT  
 CCGATTCAAGTTATAATGAGGAGGTCTGCGGCTAGGAGTCAATAATTTGATTNGGCTTATGCGGG  
 CTAAATATTAGACTGGANTCGTTTGCATCCTACTGNCGATTCTGTCGATCCATNNCANGATNNCCNN  
 TATNAAGATACTTGTNNCCCANTCGTNTCTTGNGGNCCTATTGGCTCTCTTCCCTTTGAAGGACNTTT  
 ANNGGACCCCTGTGAATTAGTAAANTTTGGCTTGGNTGGAGAACTAGCNANNTTACTTACTTATN  
 AACCAAAANCTTCNCCCTGTAGGGTTTTANTTCAAAANTGTTTNNANNCCTAACCTACNCNAAT  
 AAGNTTGATNANNTAANACCCTCGCACNCNNGGAGCTGAATTTATTCAACATGTTN

SEQ ID NO: 365 ACGCGGGACAGNCCNGNCCACAGANNGANGGCATANTAACTTATTCATTNC  
 CANGAACNCNNGNGGNAAGTATNTGNGGGATCTGGCTCAGGCCACATACANCANCTNATCGAA  
 GCTGNCGNTAAAGATGATCNCTTTTAAATTAGNGCCTATGNANNATTATAAAAATTGAANNCTCCTG  
 NGTGGAAAACAAAAANAATGTTCAAATCCATCAAATCTTNTAGAACCTTACACNCAGGNTATTTT  
 GGACNTGACNTATTTGANGATAACAAGNTANTAGATTGAAGATTTTNTCTAAAGACTTTTNTTTAC  
 NGAAATGTAANATAACTGANTNATNTTCTTTNAGTACCNGNAANTTTTGTNAAGGNNAATTTTNA  
 TGCTTCCAATATTTCNATTTGTTGGGATAACNAATTGGNATNTTTTTTANAACAAANACCTTNC  
 ATNATTTTANTAGTNTTNTTNNCCATAAANAAGTGGTCTTTAANAACCTATTGTTAAAANTCCTA  
 NCCGNAANACTTTANGGTANNTTCAACACCTGTNGCGTTATATGGAT

SEQ ID NO: 366 ACTTTTTTTTTTTTTTTTTTTTTTTTGGTGGTGTTCAAATATTTATTTTAGGTA  
 GTAAAGGCTTTTGTCTCTAAATCCTGATTACATGTCACNTGATCAANAACAATAAANGGAA  
 NTTCAGGACCNTNTNAAAAAANTNCAGCNCNTTANAANTTTCCAAAAANGGNNGGCTTCNTANT  
 TNTAACAANTTTTTGNTNANGGTNGGGAANANTTNAAGNNTTNTTTTTNACNTTTNAANNAATTT  
 TTTTTANANNTNCTGNACNANTTNTCCAAANTTNGAACTANTNANNTNTGGGNAAAAATCAAAA  
 TTTTNGAAATAAAATNAAAACNTNACCNTNTTTGGTTGGGCCCTTGGGTAACTGNACTCNAAGN  
 AAATCCTTAATTTNAAAANAACCTTCNTAANAATAACCTTTTCCCTATTTGGNTGTACCCAAAAC  
 CNATTTCCAAANACTTTTTATTCTTCNCCTAAATTT

SEQ ID NO: 367 ACATCACAACATGCTTTTTTAANNTCATTATGCATTGTGCTCACATTCCCTTAA  
 ATGTTGTTTCCAAAGGTGCTCACTCTCTANCCCAGCTGGATTCTCCGGGNAANAGGCAAAGACAGT  
 TTTGGCTAANAANAACACATGGNAANGATGGNGGNNGGGGAAAGGAAAAAGCATTNCTTCCA  
 TTTNAANGATCAAGNCNTNANTTAAAGNGTNATATTTCCCNTAGNCTTTNCNTANNGGCNTATAT  
 NTTGGNNTNNTANGAATANAANATCNACTCNTCTTNGAANCTNNTTCCCTTCATACATATTTTNTAT

SEQ ID NO: 368 ACGCGGGGCAGTTCGCCCATGGCCTCCTTGGAAAGTCAAGTCGTAGTCCTCGC  
 AGGTCTCGGCGGGAAGTGGAAATGCCCAATCCACGACAGAACAAATNTTCGGGGCTTTTACCTAN  
 CTCAACGAACCCCAAAACCTGCCGCTCATNGNGNGGCTGCTNGTGAAAAGCTTNTCCGAAAGNGG  
 AATCAACTNTNAAATTTNTNATCATNAATNANGGAAGCCCAATNGAACAAGGGATNTTNTCTNAAC

AATTGGAAAAAANCTNTNNGGTCAAANAAAAATCTTNTNANANCANNAAAAAAAGNTGGGANT  
 TNGAACTNNATNTTTTCATGGAATNAAACNTGCCACANGAACTGGTTATTATTATGGATGCNNAT  
 CTCTCCACCATNAANCAAAAGGAGGGNAATTTTGATATTGGCTCTGGAACCCGCTACAAAAAGAA  
 TGGAGGNGTNTTTGGNTNGGATTTGAAAANAAAAATTATTANTGATGGAATTTTGCCATGNTGCC  
 CAGCTNGGCACNTNCTNNGCINAAGTGATCCAACNTTTTGGCCTTCTGGGATTACANNCCGGG  
 GGCAATTTTTTACTTAAATNTTGGTGANAACCAAGANCCTTTGATTTTNANANGGAANTTTNAAAT  
 TTTCNAAAA

SEQ ID NO: 369 ACGCGGGTGGGAATGACAACTTCGGTCGTGGAGGAACTTCAGTGGGTCTGTG  
 GGTGGCTTTGGTGGGCAGCCGTGGGTGGGTGGGTGGGATATGGTGGCCAGTGGGGAATGGCTATA  
 ATGGGAATTTGTAATGAATGGAAANCATTTTGAAGGGTGGNNGAACTACAATGANTTTGGGAA  
 TTAACACAATCANTCTTTAAATTTTGGACCCATGAANGGAAGAAATTTTGAAGCCAAAACNTTT  
 GCCCCTATNGCGNGGANGGCCATACTTTNNCAAACCAACNAAACCAANGNNGGTNTTGGGTCCC  
 ACAACACCATNNCTTTGGCAATTGCAAAANATTTTNNATTTNNGAAACCAAACTTAACAAGAAAA  
 GGAAANCCNAAAAAATTGACANGGGGAACCTCNGGTACNACANATTTTNAACCTAANCCNAAC  
 NCANTNGTNGNAGGGNCTNACTTCTTAAAAAANANNTTTTAAANCAANACCTTATNTTTTT  
 TGGCAAAAAAACCCGAAGNACTGTTTTNGNANNAATTTTAAANANGTTTTTTTTTAATTTTTT  
 TTTNGGGAAANTNTAANCCTTCAAAAAAGGGTTTAANGNAAAATTTTTTTTTTNCCTCCNTTC  
 TNNNGNTTGC

SEQ ID NO: 370 ACTGCTGGGCGGCTTCTTCGCGCTCGTGGGGTTGGCCAACTCTCGGAANGA  
 AATCTCGGCTCCAGTTTCCGGAACCGGATGAATGCCCTGTTCGTGCAATTTGCTTGANGGGTTCC  
 CCCTTAANGGNTTNGCTACCNACCCAATNCCCTNAACTNNCAAATNACNTNNGGGCTTTCNGGA  
 ACTNNTNGTNGGTTGNTNNNGTTCATNGGCCANCNATNCTTNAAAAAANCNANAANCTGNTCT  
 TNANTCTGNTCATNAAGGGGGCNTTTNTNANCTTGGNANCTTNTTAAAAANTNACTTAANCACCTT  
 TTCCNACCCTTNTCTNCTTGGGTTCTCTGTTNTTAANGNCCCCCACTTTTNNNCCAAAATAAA  
 AAAGGGGNTNANACCCCTTTTNAAAAAAATTTTAATNCCNCCCGGGNGGCCCTTTANAAAGG  
 GNNAATTTCANNAACANTGGCGGNCNNTTTATTATNGGATNCCAAC

SEQ ID NO: 371 ACAATTCATCTAACTTCCGGAAGCACTTTTCAGTCCAAATGCATAAACCGTCC  
 CACATGCCCNCCAGAACCANCTTNAAAANGTCAANTTNGCTAACTTTAACCAAAGNANTTCNAGG  
 GNTGTTTTNTNAAGGCCTNNNTNATNCNNTNTTCTNNTNTAANTNATNCNCGGCCCTTGCCTG  
 NNTNCGGNAACGGTTTTTTANTTTNGCNTT

SEQ ID NO: 372 ACGTCGACCACTACAGATCCCTGGAGGAGGACCANGAACCATTGTTTCA  
 CCANAAACCTGGGAAAGGCCACANCAATTCCTTTCAAGANAACTTCGGGGCCAAACCAANAACN  
 ACNCTTNGGNGAACCCCATNGGAAAGGGGTNGGANTCAAAANAAGGANCACNAATTTCCCAN  
 NANGGCAAAACCCCGNGGATNCCNANAATGATNAAAAGGCCNTTNNGGTGANCNAAAAANAAT  
 NCNCAANGGCCCTTTTCAAAANANGNAAANCNAAGGCCCTTAAGGGANAATGNNAGGGAAAAAC  
 CGCCCTTTTTGGCNTNAAANAAGANAAGGGCCGAAAGGGCCCNCTNAAAAAANTNTTTT  
 CCTNCCTTAAAGGCGNTTTNNNAACCCCTNTAAANCCCTAGCCCNAGACCCCTNATAAGCCN  
 AATTTGGCAAAANCNATCCANGGTNTTGAAANCTTTTACNANNGNNAANGGNCNGNACCTNT  
 TGNCNAANGNTAAAAANAAGGTTTTACCCCTNTAACTCCNAAAGGGAAAGGCAAACTTTTG  
 GNTATAAAA

SEQ ID NO: 373 ACCTTGCCAGGTCTCCACCAGGCACCACAGTGGGAGGCTGGTAGTTGATGC  
 CAACCTTGAACCANTGGGGCACCAATCCACAACTGGATGCTTNCCTTGGTTTTGATGGNNGCAAT  
 GGCAACATTGANATNTTTGGGAACCANCTNACCACNGGTTTACNNGCNACAAGNCNTNTTTTAN  
 CATNGNGANGGTANATTTTANCATTTNGTTGGNTNGCTNAAANCAAGCNTTGGTGANNNTTNT  
 TNNANAAGCTNNTCGTGGNAAGNTTTNTNAAACAAAAANNANCNNGGGCNTTTTNGGCCAAAGGA  
 AATGGATTCCGGGGTNAGGTACCAAGTTTGTCTNGAATTNTNGNAAGGNAACATTTAAGGNTTCA  
 TTNAATNTTAAGGNACNATGNTNGGAGAACAAATTTGGTAATAANGGGGTAAAGGTTATGTNNGG  
 TGGCCCTTNTATTAGNGGTTTATAAANAATGNTATAANGGGCNTTTNTGTNACNNNNAAGGCAN  
 ATNNAGNGGNCAAGNGGGGTGGGTGGGTAA

SEQ ID NO: 374 CCGCGGCGAGGTACTTTTTTTTTTTTTTTTTTNTGGGAGTTTGTAAAGCAAAACA  
 TTTATAGCCATGATTCCACAGATGTNGGTGAGTAAACCAACAGCCATTCTTAAATAAAATACAAA  
 AATTCCCATCATAGTTAACAGGNTNGNAAAAANCCGGCCTTTANGGGGTAAATGCCTTA  
 AAAATTNTNCAACAANCATNGAAANNTGGGCTTTTTTTTTTTTAAACNCAATTGTTGGGTTTTACCC  
 CCCTANNGATAAAACNTNTNCANTTTNTACCNGCAGGNGGANTANCTTGGCANCAAAAAATNGCA  
 CCCCCATTCAAAANCTTNAGNANTACCTNTCCCCCTNTAACNANAAAAAACCATNTTTTTNGAA  
 TTNAAANATTTNATTTTTNGGCTNAGNTTNTAGCCNNGTTTTTTGGGATTCAAGGGTTTTACCC

AAAAAANGGAAAAATAAANCTTTTGGGAACCTTTTCCNTTTTTTTTNGATTAAANCCANGGGAACT  
NTCCANTTTCCCCCTTTAANATNCAATTTNNTAAAAANAATTCCCCCANGTGNNTNTGNCNTANA  
ANCCNCACTTTNCCCCCAACCTNTTTANNNCCCTTNATTTGAAAATTNCCNTAAGGATTAACANC  
CNAAAATGGGGGCGGNTNCTT

SEQ ID NO: 375 ACTTTGGCCTCTCTGGGATAGAAGTTATTCAGCAGGCACACAACAGAANGCA  
GTTCCAGATTTCAACTGCTCATCAGATGGCGGGGAAGATGAAAGACAGNTGGTGNNAACACANTT  
TCGTTTGATTTCACCNNTTGGTCCCTTGGCCGAACGGTCCACGGGAACACTNTNATNTTGGNTGA  
CTNTAATAAAACTGCCCACAATCTTNAGCCTGCATGCTGTTTGATGGTNAAAANTNAAANTNTTTT  
CCAAAACCCGNTTGGCACTGGAANTCGGTCAAGGGGACCCCNCGNATTCCCGGGTAAGATTGCCC  
NNNAAAATNAATAGTTTANAGGGNGGCTTGNCCCTGGTTTCTTGGTGATNCCAANCTNATNNANCT  
CNTTANTTNTTGGAGGGNGGNTAAAAANACTTTGGNNTNGAACCTNNCNATTNANAGGGNGGGC  
CNTCTTGACCANAAAAACCAAGTCAGGGGGTTTTNGGA

SEQ ID NO: 376 CGCGCGAGGTACTGCCCCATGTGCAACATTAAGATCCACGAGACACAGCCNC  
TGCTCAACCTCAAAACTGGACCGGGTCATTGCNNGGAACATCGTGTATAAGCTGGGGNCCTNGNC  
TTGGAAAGACAGTGAAANAGAAACCGGGTTTCGGGGAATTTTACCCAANTCCCGNNNGTTTNGG  
AACCGGGTNNACCCAACCCCNNTTGGGGAANGANCCNANTACTTGNCCAACCTTCNGGCTTNCCC  
TTTCANNAAGNTTGANCACTNNTAAGGCCANTACTTTGGCTTATGATGANCAAGTTNAANCT  
GTNCCCNCTGGAACCGG

SEQ ID NO: 377 CGTNCCTAAANTGAGTATCAACTGNTNNTGCCATANCACTGTGNNAANTGCA  
ATCCTNTAAAAGGCGAANGTGTNACTGAGGAGCTTGNCAAAGTGAAGCNTTCATCCANCATATGG  
AANNTGTGTCNGCNATATTANTCACNANGTTANTGGCTGCTNCAANCTTNATNACANTANCTTGNNTA  
ANCTTNGACTGCTTCCTNANGAAANNCCCTGCNNNGNCATNTCTNTAAAANGTNCCTANTNNACAN  
TCGACTTTATCTCCTG

SEQ ID NO: 378 ACAGNTGGACCTNNTGNNATTAGAGGCNNTNNNGGTCACCAAGGCCCTNCTG  
GCCCCACTGGTCTATTTGNCCTCCTGGACCTCCAGGTGTAATNCGGTNAACTNCTNGACTNAAGC  
AATCATGGACNTTATNCTGGNTGTACCAGGGGAATANCCAACCTCTTTACTGTATGCANANNTGT  
ATGAAAAAATAATATNCCTTTNATGACTGGNNAAGGNTTAAANNTGTATCNTCTGAAANG  
NAANGNTTANNTAGGNGGNTTTNACTNATTCAAATANTTAA

SEQ ID NO: 379 CATGCCTCGAGCGGCCCGCCATTGTGCGATTGATTTINGCATTNTTTCNNACCT  
TCTTCAGCTTCANCCCTAACAATGTACCNTATTCTNCAATNTAAAANTTCCNGGGCNTCCNCTNCNT  
NNCNTCTNNGCTCTCNGGTCCGTCTCTCCTCTCTTTTCGATN

SEQ ID NO: 380 GTACTCTTGGTTTGTCAATGGGACTTTCCAGCAATCCACCCAAGANCTCTTTA  
TCCCCAACATCACTGTNAAATAAATAGNNGATCCTATACCGTGCCAAGCCCANCTNACTCCAGAC  
ACTGGGCCTCANTANGACCACCGTCCNATACNATCACATNTCTATTANTAAGCCCNCCCTAAA  
CCACTTNCNTCATCATNCAAACANCTTNAAACCCCNNTGTGAGANATNGANGATGCTNTANCCNT  
TAAGTGGTAANCTTGANATTNAAATAACACAACCCNACCTGTGGGNGGNTNAAATNAAATNANANNT  
NTCCCANTTAANTNCCAANNITTCATNTNNCCANATTGANNANNGGAGCCTTTATTATANATNA  
ATTGTACTIONAAGNTCGGANTATTGACCCTNCGTANATTTNANTNTA

SEQ ID NO: 381 ACTCCTTGGCCGCTCACTAGCACTCTCCGCTGCTTTTTAAAGGCTTCATTG  
GAGCCAGCAGCGTGGCCTGCTGCNAAATGAGAGTCACCAGNCGTTTAAANCAGGAAGGACAGCNC  
CANGGAAANCCACCANTGTAAANNTNCNTTTGGGCCCGGAAAACCTTATTNNGAAATNTCTCNT  
TGCCCCCGGATTTGNTCTNGAAGNTNAACTTTTTCGGNAAAAATTATTATNNTTCCAATTTTCNCC  
CCCGGATTTGATTANCAACCTCCCNANGGTTTANCAATTNANAACCACNAAAAAGGNGTTTGCCA  
TTNGANACTTNANAAACTTCANCANCCCGGACTTGAAAAATTTTNGCCNTTTTTTNGGGAAAAA  
ATGAANGGGAATTNCCGAAAAGCCACNCAACAAAAACCTTTCCCATNAAAGAAAGGTNCAACTTC  
ANATCCAATNCAAAAAAT

SEQ ID NO: 382 ACCGTTGCACTCCAGCCTGNNCTAGATANTNAGATNCTGTCTCAAANNAATT  
AATNNTNATTTANTAATCATAAAANTCTTGATCTANGCCTACNATNTTANATGCATNTTNTNATN  
TNTNTANANNAATATTAAATTAATTNATCTNTGNTACNANTTCTATAATTAATTAANATTTNTN  
TAATTNAACTCTATNTNTNTTTNATATTINCAATNTNACCTNNAAANCTANNAAATTNNATNTA  
TATNAATTTCNTTAAATNTAANATNTAATTTNTATANCNAAANTAANAAATTNTAACTNATATAA  
TANTC



SEQ ID NO: 383 CGTACTACCGANATGCCCGTTCTTACAACCGGTNTCAAATCGNACACTGTCA  
CCGAAAAAGGTGTTGAAATNGAGGGACCATTGTCTACNTNAAACCAACTGGGATNTCGNCCANAT  
NAACANTGGNCTNTNNAATGAAATACAGATGCNTCTTTGAACTGANAGTNGCCCTACTNNTAA  
CTATNAACTNGGCT

SEQ ID NO: 384 ACGCGGGGAATGTCTGAAAGTCCATGACCTGTCTTTAATAGCGGATTATNAA  
ATTGCNNCCAAAGNAACAAGANTTTTTCTNTTGAACCTTNATNCCATGNATCATCTGATNTCATTCA  
TGCANATTGGNAACATTATATCANAGTTGTGNTTNNNTAATTATANATATATTTNCTTATNAAATT  
TTAANGTTATNTTCNTATAATNAANTNNTAATNNTTATCNNTAATTTTTTTTCATNTANTATATCTA  
ATTTCTNTTT

SEQ ID NO: 385 ACGCGGGGGACTGGAGACACTGAAGAAGGCAGAGGCCCTTATAGTCTTGTT  
GCCAAACAGATTTGCAANATCAANGANAACCCANTGAGTTTCANANAACCGCTAANTANGTTATA  
GANATTCTAGTNCATATCATACATATTTAAGTATNNANTAATTNNTTNTAATATANACNT

SEQ ID NO: 386 ACTTTTTTTTTTTTTTTTTTTTTTTTGNCTCTTTATTTTCAANNNTTGGCTTANTA  
ANATTTTTCTTATTTTATAANGCNATTACNACAATTTANGNAACNAAACNATTTNAAACNAAAAAAT  
NNTAAATNGTTTTTTTTTAAAAAATAAAGTTGNTGNTTGCAAAAAANTCCNTTTNATNTTATCCCCC  
CCAAATNTNATTTTATTCTTTGCNCTAAACCAAAATANCTTATNGAAAATTNNTTTTAAATNNTCTT  
AACNCNAAAAACCTTCCGCTTTTAAATNANNTNAAATNCTNNTTAACTTTNATNGGATGCNTAAAC  
AAANCNAACTTTNATNCAATTTNAAATNAACTTCATTNATTGGACNAAATNCC

SEQ ID NO: 387 ACTGCNNGGACTTCTCCTTGCTGCTGCCATGTGAAGAAGGATGTGTTTGCTTC  
TTATTTTNCNATAATTGCNNANNTAGGCTNCTTTCATTTTACTAANACANTGTTGNATTTANAAA  
NACTTATANNCCANTNTTTAANACANNAGTNTTTTAAGTA

SEQ ID NO: 388 TCCNCGAACCNGCCCGCCAGTGTTGATGGATATCCTGCANAATTCCAGTCTCT  
ACTGGCNGCCGTTACTACTNGATCCNAGNGGGNTNNCATTTTNTTCCACTNATATATGGNNATACT  
TANCTATTCCNNCTATNNTATTTNGANCNNCNTAATTATTTTCANATNATTATTAACCCANANTN  
CNATGTATTTNTTAGTTANATTTNACCTTTAATATTTNTNTANNTTATAAANAATAATTCANTCNA  
NNTNAAATTTNAACTCTNTNATTTATNTTATAATNTAATGANNNATTTNATNNNTNATATTATNA  
ANTATAATACNANCTATAAAANCNNCTAAATNATAATNNAANNCANATNTANCANTNNANAATTT  
CTNTGAAT

SEQ ID NO: 389 ACTTTGTGGATAAGAAAAATGGAGGAACACATCTNATGGANAGTGGGCATTG  
ACATTNTGGAACAGGTAACCANCATGTATAATNAAATTTATAAGTTTCTTTTAAATATCTNAAATNG  
CTCTNNCAGATATNTTTNATNTNANTNTNNTTATTTATNTTTTCTTCTTTAATNTTTTAAATATAAAA  
TNTAATNNATTTCTTGTAATTANNTNTATTTNTTTTAAATTATATTANNAATTATTCTATNAAANAT  
TTNATTTTNNANTAATATANTTTNTTNTANATTAATTTTATTANAAATTTNNANNATTTNTATTA  
NATNGAGNATTTATTTNTTTNTTANTTANTNTTAAANTNTTCTCTTATNTNTAATTTATTATT  
TTATANNAATANTTCTTTNTTTTATATTTTANTNTNATAATTTNTNAAATNTNANTATTTTNTNAA  
TTANTATNATTTTATNAAATNTTTANTNTANTANTANTNTAATATTTAATTTNATNAAANNNTAC  
ATTTTANNNNTTTAATANNTAATANANTATTAATTTANNTTTTNTNTNTATTTANATNTATTTT  
TTNTNANTAGTATTTTNTATTAATTTATTTTNTAACTGATATTTTNTNTATTT

SEQ ID NO: 390 ACCNNGGGNNCNCNTATATCGCNAGNCTNTTCCCTNCCNAAGAAGCACACTTTG  
TGANANCCAATGGGAAGGAGCCTNANCTGCTGNAACCTATNCCNTATNAAATTCATGGCATAATAG  
GTGTTAAAAANAAAAAANTAAAGGACCTNTGGGCTAAAAANNAANATAANTGGNAAGANNCCAG  
NGGAATCNNTNTAANNTTTTAAAAANCAAATGNCCTGCTNAACANTNTTACANATTTTCATNAG  
CACTGNAGGTTTGTGTTGNACCNCINGNNCTTATTGGCNCNATTAGTTGTGTNTTTGNAAACATAA  
ACNNTNTNGACTCGATAANCATTTACAAAAAATCTTGGNNTTGAGCCATATCNGNAAAAAAGGGAT  
TTCATNTGNATGTANAAGATTATCTTTAAGNANTTCTAATTTTGGCAGTGAAGTCAAAGGCTGT  
NTG

SEQ ID NO: 391 CAANTNTAAAGACCCTNAGGAGTTCATGGANCACATATATGTTCCCANAGGA  
CAGTATNTANCACCTGCGGCGCCGGCTNNGTCACTTGNGGATGGTGTGAAAAATGAGAATNNTTA  
TATACAGACTNCTGAACANTGTNATGTNTNTTAAAGGANGATGNNTCCATTTNCNTTGGNGAAT  
AAGAACATAANTTTCTAGGAGAACTAAGGTTGATNAAGTNTACATGCGGGAAAAACAAATTNATN  
TGCTTGTGCATCATGNCACAAAACTNNAGNANAGGCCACCTGTGAGAGACATGTCCATGATG  
GAGGTTCAACATCTCAACATGGNAATGACCTTTNGTCAGAATNCNCCATTGGAAAAATNTGATTTT  
GCTGNTGNTNNACATCNTNGTGCTTTTTNCCTNAGGGACCANATGAACTCCAAATGCTTGGATGT



NTGATTCANGAACTGGATTGCCTTANTGGAACTNGANGAAANGGNTATNCACCANTCTNTACCT  
GANA CTATANAGAAAACGTTCTNTTTNATTGTNTTCCCAGATGTGTCA

SEQ ID NO: 392 TTTTCCGAAGCCGGNCCGCCCCGGGGCAAGGGTACACCTTGGTTTGGGTGTTT  
AATAATGGGGNGAATGGGGGTNCTCCNGGNAACTINTGCTTTATTATTTANNCTTATCAATACGC  
CCCAANCATTANATNANTCCATAAACTACCNCANTTANAANNCTCCCCCTACTCACACTATCCT  
CCTNCTCGGNTGCNAGCCATTCCACACCNTACCANTCCACAANCTCNNCCATNACTCCCCCTTN  
CTACCTNCCCATCANCCCCCTCTCA

SEQ ID NO: 393 ACGCGGGGAGTTCGGTCGCAGCCGGGATTTGGGTGCGCAAGTTCTTGTTGTG  
GATTGCTGGGAATCGTACTTTGACAAATGCCAACTTTCGGTGAAAACTCTTGACTGGGTAAAG  
AANCCATTNCCCCTTGANGNTTGAACCCCANNTGGANCANCCATTTCGGANGAATTGTTCAAGGGN  
CAAAAGATTCCAAANNNTAAGGGAAAGGGCATTCCCNCTCCTNGAACCANCNANAANGCTTNATN  
CTTTTNNTTGAAAAACNANCTGGGAANAANTGGCCGCCCCCTTGTTTNTNATNTCCAAACNTCNN  
AAAAANAAGTNCNCCCCTTNGCANCTTNGNTGNNCTCCGTCTNTAAAAAGGGTGGA

SEQ ID NO: 394 CGTGCACTCTTGNTTTGTCAATGGGNCTTTNCANCAATCCACCCTNGANCTCT  
TTATCCCCGCATNNCTGNGAATAATTCTGGANCCNANNCTNGCCNNGCCCATNNGTCNCACACTG  
GNCTNATTNNGACCACCNACGNNCNTTGACANTCCTTGAGCCTCCCAAACCCTNCATCTCTTG  
ACATCTCCAACCCCTTGNGGATG

SEQ ID NO: 395 ACGCGGGTTGAAAAAGAAACAAAGGAATACTTTGAGAGTTGGGGAGAAAGT  
GGAGAAAAAATGTGTTTGAAGCTCTTTCTGAGCTCATAATTTTAACAGCTAGCCATTGTTTGCAT  
GGAAAGGANNTCANAAAGTCAACTCNATGAAAAGGTCNACATGCTGTTTGCANATTTGCGATGGAT  
GTTCCCCCCCCNCCNCTNCCNACNCTGNGNCCATTGACNGGTGCTTCTCTTGCTGATATACTGGC  
CNAGANGATTACATCNCCTGNANTTATNAAAGGCTCANCAGAGAACNANNCTNTCANGGCTGCTNA  
NGGAAGCNAANAAGGCTNACCCANGCNTCTAAANAANACNGCANTGGCTGCCGCTTAAGGCCCC  
TACTAGCGCCCTCCTANCNANANANGGGNNNCTGGGAAAGTTCATCTCCCGNNTGGTGGA  
CCGCTTAACCTTGCTATTCTNTTCTATTTAT

SEQ ID NO: 396 ATNCNACCCTTCAANTTCATANGGTGGGGGAAAACCCGCGNGGACTTTTANAT  
ANTCCNGGCGGTTTNCCTGGGAAACNTCCTTGTCNCTTCTTCTTGTTCAAACCNNTNCCTTTNAC  
CGGATACCTNTCCCCCTT

SEQ ID NO: 397 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNTAGCACTTTTTTTTAATCTGGTAANAA  
CAGCATNANAAAAAACTTCANTNTAAAGGAGAACCAANAGCAGGGGGATTGGCAGCCAACTA  
TTACCAAGGCNAANCTTGTTACAAAAGGAACCTTTGGGNNATACAAACNNTAGGGCTATAAANTG  
NNACTTANTAAAGGGNTGTGTNGGGAANTTNAATTTTNTNATNNCAAAGTTNANTAAATNTANG  
TTTTATNTNATGAT

SEQ ID NO: 398 ACGGGGTCCCCTACCAGACATTGAATCTGCCAGTTCCTTGATCTTGAACCTCT  
CAGGCTCCANNACTGTGAGAAATATGTTTNTATGATATTTGAAANCCNCCCAATATATNATCNGTG  
NTATNNTCANNTNANNTANGACTNNGATAACTATTGAATTTTAAANAATTAAGNTTTCATATT

SEQ ID NO: 399 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGGACAGTGCAANAANANA  
GGGGTGACCTGTGAATTGGTGCTGGGGANCTGCTNAGGCCCAATGTGAGGCANCACTAAANANAT  
GANTAAATTTAGGGNGATCTTTANCTNTCCTACCCAGGCAANAAGGGTTGGGGAGCGGGGGTNC  
CNCAANTTGGCTTNCANNGTATATTCAAGCCATTGGGCTNNAANANGCTNTNTNTNTTTTGCTC  
ANATTTACNNNGNTTGCTCTGGGGCCANAAAACCTCACNAAACCCNNTGTNATTTNAAAGGANNA  
AGGCCCTTAAACTTNTGTTANCCANCANANAGNTCCATATNANCANAACCTTTAATCCTTGGA  
NAAA

SEQ ID NO: 400 ACTTCACACANGATCCCAACCCCCACNNANNTCAATGTGACCNCTCTGATC  
ANTCAGCTTCATTGNCTGAAACAGCNTCTCTCAGGANATTCAATNATANATANTCTTTTGTAAAC  
CATAANGANAGCTATTATNTTNTCTNNCATNANNACTATATTATTTATTTTNA

SEQ ID NO: 401 ACGCGGGGAGCACGGTTCGTTTTTCTTTANTCAGGAAGGACNTTGGTNTTTA  
GGTTTACANACTTTTCAANNTAATNANTTATTTTTATANTGTTATGATTTAAATTNNAATTNAATA  
TCNCNTNATNTGTTTAANTTTGTNTGTTTANTNTATTANATANTTTTTATGTTTGTNTATNATATT  
NATTATAAATTTT

SEQ ID NO: 402 ACTGCATTTTTTTTTTTTTTTNATAAGGCTTATAACTATGGCTGGATCTTTTG  
CTCTANTCTTCTAAGAAGGGCNATTTTTATTTTNTATAANTCNCCTNTAATGTCATNTGGANTATTA  
ANTTTGNANATNNTTANGCATTANGTANANTNTTATA

SEQ ID NO: 403 ACGGATGCTACTTGTCCAATGATGGTAAAAGGGTAGCTTACTGGTTGTCCTCC  
GATTCANGTTANAATGAGGAGGTCTNNTCTNTNTATATTTATTNNNGACTTTATANNNCTATTTT  
TTCTTTNTTTTAAATACAAATATTTNTTACTTANATNNTTTGAANTAATTTCTTANATAATTGTN  
NGCTNTTAATNTATNINTATTATNAATCTTTATNATTAATTTATTCTAAATNNNTATANTNTTAN  
TANNNGGINTNCANTTNTATNANTNGTANTATANNATTNAAATNNANTAAATTATAAATANTTTA  
TCTATTTTNGTNTTTTATAATATNTTTTAAATATATTNANTTTATAATNTNCCAAATATNTTATTCA  
TTANITATATTTTTTNTTTTNTTNTNNTAANTTTTNTTNNNTATANATAATNCTTNCATATTTNAANAT  
ATNTNTATNNTNGATNTTANTNTNTATNTT

SEQ ID NO: 404 ACTTCAAGATTAGGANNGTGGGTTTACATAAATGTATTCTCTGGTNANGGT  
GGCTGGNATATACCTTGACCCACCATCTTCAGAANGACCCATGTCACGTCTTGACCATTGGGANCA  
AAGCCATGTTTCACACTGACCTAATNCAAGAGTNTGGAAGCATTGGGCTGGTTATACATTCTATTTT  
TTAAATTNATCCTNCCCCCTNTNAGGCATNGANAACCTTTNATCANATCATTTAAAANTGGCCTNT  
NATTTNTCNTCAANATANTTATNGGNCNTNTCTTTATNATAACTCTNGTATTGAAAAATNCTTNC  
CTTTCCNTAATGGGGGANNCTNATTTTNGCNATTCCTTCNTTNCCTTTATNTNCAAAATTACCNA  
ATTNNCACTATT

SEQ ID NO: 405 ACNCGGGGGTGAAGNGTACAAGCTCCTCCTGTTCCACCCCTGAATTAACCC  
CTCAAACACANAACCNCTCCGNTGCTTGTTCACAAGNGGTGTCAAGGCATCAATAACCTGCANN  
ACTGTGNGAAAACTANTAAAAGCAAATNCAATGTTCATGCTGNAATCCCACTTTGAAAAANTATTG  
ATAANATCTACTTGACTCTGCTTCAACANGCTGCTCGCNCCATNTTGNACCACAACATAACCAAC  
TNCATNAACANCCAAAAACAACNTNNTAATAAACTATNATAACATNAACCNCTCCNANTTCATTN  
TNNGATCATTANAAAGCCTTNCNTTTTACCTNATTCATCNTACATNATTTATTGCCTANAATNTTT  
TNNTTNTATTGGGATNTCCACCANCCAGNAAANTNCTTTNNCCNNTNAAANTCAAAANTNANNT  
CTATTNNNTTCCAAAA

SEQ ID NO: 406 ACNCGGGTGATCCTAATGTGGNTANTACTNNTNTGGANANNACTCCCTATAN  
GGCTAACNNNCTCNANCANGACANANNATAANGAACAAAANATATTCCTCCTATGTTACANGCTT  
NGAAGGGCNTACAGGATCTACGAACCTTGGAAGAAAANCATGGNGCTCAGGAATTCATCNTCTAA  
CATTTCACTTCCCCACCCACCCCTTAATGCTCCCACTTTNGCAATNATCTCTCTTTGGNTTTTAAAN  
ANAAAGGGGGAANTNGNGCCTTTGTTTTNCAGNTNTGCAACAACACANCTTTNNGCNTCTCAANC  
ANNANTNGGAGAACTCTAANACAAAAATAATTTTTTNTATNAAAATNACGGGTTGTATACANT  
ACTTNTCATGGCCNTTACAANGGNCCTTANNGGTTTTTTTNTANCNTACAACATTTCCCCCNA  
ATTCANTNAAANANGNTTNTTNNCCANTNGAAANAAAAANCTTTNNCTTATTNTANAANTNTTAA  
AATNCNATATGGCTNAT

SEQ ID NO: 407 GCNTTTGAAAGATTTGGGGCCCCTTCTANGATTGCATTTCTTCGANGCCGGCC  
CGCNCAGTGGTGGATGGNATATCTTGCAATAAATNNNNNTCTTANCGTNNCTTCTGCNTTATACT  
TCTCTTNGANTTACNTTTTNAAGNGNTTTTCTTATTANCTAAATTNNTCTAAATCNATTATNATAAN  
CTGNNTCTCANTTNCCTCAANACATNTTCTAA

SEQ ID NO: 408 CGTACCCATCTCAGATGAATGGNTACGGATCATCACCTACCTTTTCCCAGACG  
GACAGAGAACATGGTTCANANNCAAGTGCTAGGCCCTTATGANCAAANGANGAATNATGCTCNG  
GACAGTGTAANNATATGTTTATTATCNCNAAACGGGANGANNANTTAACCTACCNNTTCTANAT  
CTGAAAATGNNAAAAAANTCTCGATGTTTAAAAANACCCANGGGCTTGTNTTTTTTTTCAAAAAAN  
TNNATTNNACNAATNTNTTATTGGGGGNNTTCAAACCTTTTCTTTTGGCCCAAAAAATCNCCAC  
CTTTT

SEQ ID NO: 409 ACCCGGNNGGGTNGCTCNTATNAAAACCTCATNACAAGNCATTTATTTCTGT  
AAACTGCCATACAACCTTACATGTTNANTATTAATTGCACNAAGTNTATCAANACTTATTNAANAA  
CACAAAATTAANTNCTATTTTAAACGGTTCNTCTTTTCATTATTTAANTACNTNTNCTNTNTTANCTT  
GNACATTTTGTTAATANTNANTTTTAT

SEQ ID NO: 410 GNNCGNCCGGGCACGAACCNCGGGGANAGATNNANAATNATTGCCAGNC  
TAGATTACCAGGTGGTAACCCCTANACCNTCTGCATAATGAANTANCTATAAATAACTCTTCANNG  
GGAGCCANAGCTAACGACCCCTAAACCANACGAGCTACCTAAGAACAACCTCNGANTNGCACNNN  
NNGGTATGTAANNGANNANTGGNAANANNTATATGTANAGGCNACNAACCTACCGCTGCTGGTG

ATATACNGTTGTGGATACATAGAATGTTNTCTTANCTNAAAAAGTCGGCCNCNGAACNNTGTNNNT  
 TCNNTNNTGGNGTTGACNTGTTATTCCAAATGAGGAACNGGTTTTTGGNCACTATAGNAANNCCA  
 TGTTANACCGTGTGNTTAANGAATNGTGCTNCAAAATTTTTCTCCCNNTTNCNGNACCCCCCTTAA  
 GGCTAATTTNCANNAC'INNNGGGCGGNCNANCTGGTNACCTTCATTGGGACCCCCCNNGGG  
 ANAATTANCNGCNTTCCGTGTTNCNTNTTGNANNTTTTTCTNGTNAAAATTTCCCNCCNNGNCC  
 CGANGTTTACCTNC

SEQ ID NO: 411 ACCTAACAAACCCACAGGTCCTAAACTACCAAACCTGCATTAAAAATTTTCGG  
 NTGGGGCGACCTCGGANCAAGAACCCNANCTCCGAGCAGTACTNNGTTTTTTTACNNTTNCITTTTT  
 TTTNAAGGTAGTGGGCGTGGAGCNTGAACGCTTTNTTAATTGGTGNCNNTTGAATGCCTACTNTG  
 GGTGTTTTAGAGCTTNCITTTATGTATTTCATGGTTCATTGGATNTTTTTCTTNTGNNTTNTCAGCTTAN  
 GTTACTNGTTTTNGAATATATTTTACNNGCATATNTGGTTATTNCTTGTACATATGNTGNCAACN  
 TTNAATAAANCCCTTNTATCATTATTTGNCATTATATGANAAAANGTGTCTTGTNTGGGCTTTTTAT  
 TTTCCNTGNTTTATTGNTGNCTNACCTTGTGNTNTCATTTTTTNNGATTTTTTTTTTATTTTCAGCC  
 AANACTTNTNCNCTATACCCTTTTTTNTNGTNNCCNTNNTTCCACGTNATTTAANATTTTTTAAA  
 GGGTTATGTCTTCTNCTTGGGGAATTNTTNTATAANCTTTTTNTCTTNTNTNGGGNTCCTAANCTAN  
 NTTGTGNGTCCCCAANCTCTTGGGAGGNNTTTCGTGGCG

SEQ ID NO: 412 NTGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGAGAGGAAAACCCGGTAATGA  
 TGTCGGGGTGTAGGGATAGGAGGAGAATGGGGGATNGGTGTATGAACNTGAGGGTGTNTTNTNT  
 GTTAATNAGGCTTTTATTTTGTCTNNTNTTGTGATTTATNTATNTNTNTTNTTNTTNTTATNATT  
 TTNNTTCTNNNTCTNTNATTCCTNTTTTTATTTTTATNTNTTTTATNTNTGTTTTNATTGTATCTT  
 NTATCNTATTTCTTNTACTNTTCTTCTTNTGTTATTTTTTATCTTTTNTCTNNTNNTGTCTTTTN  
 NNTCTNTTTTTNTTCTTTATATTTTTTATTTNTTCTTTTATNTTCTNTCTTTTCTTTCTTTCTCNT  
 NTNTATTTTTNTTTATNACTATNTTCTTTTTCTTTTTTTTTTATTINATNTTCTTCTTTATNT  
 TTTTTTTINATATNCTAANNTATTTTGTFTTCTATNTTINATATCNTNTATTTTTTTTTNTATNTAT  
 TTTTTNATNTNTCTTNTTCTTNGTATTTTTNNNTTTTTATTNATATTTTNTATCTNNNATTATNT  
 NGTTTTTANTTCTTCTTCTTTT

SEQ ID NO: 413 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTCCCTCCCCACANAACCCATCTCAAAT  
 CATTCTGTTAACCACCATTCCAACAGGNCGAGGAGAGCTTAAACACCNTATTCCTNNGCCTTGATT  
 CTCTNCTATTTTTAAAAGGTTNCNANCAAAANTANANANANCNNNTACTTTGNGGACTAACTGNT  
 AAATNTACAAGGGGATTTACANGGTNCCGTCGGCCCNNTAANGNAGAACTAAGAATCNTNTNTT  
 GGACAACCAAGNATCNCCAACCTCNGGAGGNTTNNCNCCTTACCTATAAAANTTCCCNATNNNG  
 TTACANTAAACCCGGNGTNCCCNNTTACCTGNTCNTAAGAAAACCCCTCTNNTTCCGGGGNCCTN  
 ANNTTGGNTCTCCTTNCAAANGTAATTTCCAANNNTANNCATTTTCCCAAATNGNAANNNGNGG  
 TNAANNCCCTNCCNTNANTAANNCCNNGNATAAAATATTCNAANATTTCCCNATAGNGGGCCCCNG  
 GNCCCNANCCNCTTAAGGGCNAAAATCCCAANNCANNGGGGGGNCCTTATTTTTNNNNNNCCN  
 AACNCNTGACNNNACCTTNGTNTAAATTNNTGNNN

SEQ ID NO: 414 TCNAGCGGCCGCNCNGNCGGGCNCITTTANANATTGGCNCTCATTGGNTGANC  
 GCGAGACNCTTAGNGCCATGAAATNATTACACAGTAGGAACAGGTTCTTTATNGAGACTGGACTG  
 CATNGCTCTNAGATGATGINTGNGGGTAATCNCAGACTCTAACACTTACTGTATTACTTTATATG  
 CGTTAGTTAACNNTCTAAGNNCTAAAANACGTATAATGTTNNNTAGGACGCTCTTCCATNTANTCT  
 TGGAATACCCANNTNNTGGNNTACTATNGATNCTAAGGGGTCAGGCTTATGGCATGCCTTATNGC  
 TAGNGCTNAANGAGCGTAAAAATGGCATCTNTNAAAATCTCTATGCATAATCTATGCATGCCAATG  
 NNAGCACAAATTTCTNTCNAAGAAAAGCCNTTCAACATGAAGNGNACTGCNTGGANGGGAAGAAC  
 ATTGGTACTTTTACTTTAAGANANC GTTAATTTTTCNCAAANGNACATGGNANAAATTGATTTTNT  
 TGNAACATAATATTGCGCTTTCNAAGNTTATCAAGNGGCANNNAATTTGTTTTANGTCTCGGN  
 CNNGGCACCCTANGNGANNCCNGNCCNTNNGGGGCGTACNNGGGACCANNTTGGTCCAATTTG  
 GGNAAAATGGCCNGGTGGNCCCGGGGAAANNTTCGGTCNGTTCNNAATCTTCGNCNGGACNTNA  
 GNTANG

SEQ ID NO: 415 GGTACTTT  
 TTTCTTCTTCCCGNGTNCNCGGGTNTCATTGNAATATAGNGACCTGAGTCCAAAGCAAACCCAC  
 NTNNTTTANAANAGNGGGGNTGANTTAAATGTTNTCTNTGNGAAAAAATACCAGGCCTTTAT  
 GGGGNTNCAAAANTNTCCTAACTCCNGGGGCATAATNTGGANAAATATNTTATGTTCCATGCNT  
 GGNCCCNAAANTAAAGGAAGTGAACCCNAGGGCTNANAANCTGTTNTNATTTNAAAACAATTGA  
 CTGNCCATTTTTNTAGGAAAACNGGGANTNCTTTTTTGCCCCNAAAAAAGGTNATTGAANGAN  
 ANNTTCCNGGGCNGGCTTTTNAANGGNGNATTTCANACANTGGNGGGCGTTANTTTNGANTCC  
 NATCNTTGTCCAATCTTNGNGNAAAAATGGNAAAAATNTTCTTNGGGAAAAATGTTTTNCNNTNN  
 ANTTCCCNATAATTTNANNCGAANTTNAATGTAAACCGNGGGNCCAATATGTNANNNNANTTNT

TTANTGGNNNNNTNNNATGTCNCTTTTNAATGAAANCTTNNNNCCCCNNNTTTAAAAANNNCCACTN  
CCGNANANNGNTTNTANTNGGCNTTCCTCCCTC

SEQ ID NO: 416 TCGCGGGTTCTCGTGAGATCTGGTNGTTTAAACGTTTTTGGCACCTCCCCACA  
CTGCTTGGATCTGCTCCTGTCATGTAAGGCACCTGCTCCCAGTTTGCCTTCTGCCATGNGGAAAAG  
TTACCTGAGGCCTCTTCAGANGCAGAAGCCGTCATGCTCCCTGTACAGCCTGCTTAACTATGAGCC  
AATTNAGCCTCTTTCCCGTATAAATTACCCNGTCTCANGTATTTCTTTATTCAATGCNTGANCAAGAC  
NACTACTAAGACTTTATTGAANAAATTATAGTTGCAAAGACAATATTGCTTGTGNNCANGGATAC  
AAAATTGAGCCNTTTGAGAANAAACCGAACTTCAGNCATTANACTNGNGTANTTTGNGGAATC  
NACTGATGTTTNTTATAAANNCCCTANTGGTGGAAGTCATACNNNNTCNAAGTAGNTNNGTTTNGG  
GGTGANATTTAACTGNATNCNCNTTGGGGTANAAAANGGTTTTANNGGACAAAANGCATTATAT  
NNTTAGGNGNTNGGANANNANNGCCTACNTAAAATCCGGAAGCNCCTTGCCNCTNTNNNTN  
TTANGGNAANTTTGGNGNATCGNGNGNGNNTTTTTTTGTGTNCCGAATTNTANNAANTTTNG

SEQ ID NO: 417 TTGAANCTTCCTTTTGGGCCTTTCTTTTCCNACCTTGCCGTTTACCGGATACCT  
GCCGCTTTTCTTTTNGGAANGNGGGGCTTTNTATAACTAACCTNNAGGGNTTTAATTNNGGGGG  
NGNCTTCCTCCAAACNNGG

SEQ ID NO: 418 ACAACTAANATTTTATTAGNNATCGCTCNGCTTACACACTCCANGCAGGAAG  
TTATTTAAATCACCTCANANAAAACCTGNGTGACCTAACCNANTACNTNATATGCAGATCATGAN  
TACTTCCANANTANANTCACCNCAAGTNNNANGGNCNTAGGCCCTACGTGNANGANATGTNGCT  
TANNGTCTCANTGTAGGGANANTAAANAGCGTGTCTAGCTCNCNTNTCTACTNNGACAACNCIT  
ACTNNCATGATGANTCGTACATANTANTCCGTGCGNCACAAACATNGCTGANGNANGTGGCTTTA  
CATTTCTTNCACACGATNTCTGCGACCCGCCNGGGATTAATACTCCNGTATGCTTANNNTAAAG  
CTCCACAGCTANATTAACAAAATTGCNNGGNAGANCANTACAANNCTTCNGNTTACAANTTTAAG  
GACCTNGGCGGGGCTTNTATTCTCTACAGGNANCNNGNCNTNTATTAAATAAACNCCGATNC  
ACCTTACCNNCTTTGTTTNNCCTTTTTCCGGCTTNTTTGTCNAANCCTGNTGAANGCTCCAAGTTA  
CNCTTTNCTTTCCNNGGCGGGCGGTGAAAGGGCGANTTTAAACNNTTGGNGGNGGTTTTTTGGGTC  
CCNNCNTGNTCNACCNTGGAGAANAAANTNNAANTGNTTCNNNGNGAATGTTTNC

SEQ ID NO: 419 GCGTGGACGCGGCCGAGGTACTAGAACGGGACTCATCCAGAAGTACTATGCC  
CTCCTNNGTTAANGCTCAATCATTTAAGAGTAAACNCAAGGAGAAGCCNTTCTGTATTCATACCCN  
TCTATGTGGATTAAAANGANCNNNANTGAGANTGGNCTATGATCTNCCNNGGCTGTGCTNCTCTT  
CTACNANTCCTGCTGATCCGGAACANTACTATGNCNATAAAAAGNNTGNGAGAAAAGNNCNCGNC  
TAGNCAATCTATAANGCGGNNAACANCTGCATACNNNAATGCTGGCACNNATNAANTGGCACT  
NAGTNGGTGACCTATGCTTATTNTATGCATGATTGNGTAGATTTGGATTNNGACGNNAATCAANG  
AAGAATTNTAGNATCTNCCGTATCTNATATTNAGAAAGAAAAAATGGTGCTTCAACNTAACAAAT  
AATGAGTNTCTCGGCCCGCGACNACCCTAATNGGCGATTTCCAGGNCACTGGCCGGCCNGTGCCT  
ANTNGATTCCCAACCTTGGGNCCNACCTTGGNGGAATCATGGCNNNAGCCNGTTTCTNCTGGGTGA  
AATTGGTATTTCGNCNNNTTTCNNCAANAATANNANNCNGNTACCATTAGGGTAAACCTGNGGG  
NCCAAAGAAGNCTTAACNCNATTATTNGGTGGTTCNNNCCCCTTTAATTGGGAACNTTTNGCCC  
NTTTTTANTAANCCC

SEQ ID NO: 420 ACTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGGAAGGGGGNGTNNACCTNNANCCCTT  
TTTNAATGGGGGGNNGNTTTAAGNCCNACNNNGGGGTAAAAATTTTANCTNTNTNANANGGTT  
TTTNCNNAGGGCCNAAAAACCTGTNCNTTTTNGGACNAACAGTAAANTTTANANGGGGNTTNA  
GGGTTNTGGGGNCAAAATTNAAAGTTAANNATAAAATTTTNTTNGNANAACCANTTTNACCAGCN  
TNGGANGGTTNGCCCCCTNTNCCTATAAATNTTCCNANTNTTTNGNTANANAAACGGGGGGGCTN  
TTTNAACTGTTNTAAGGAACCTCGNNNGNTTTCGGGGGTTAANNITTTGGNTNTCNTGCAAAGTT  
TTTNTAGTTAATTCTTTNTNCAAAAGGGATAGGGGTAACCCTNNCTANATTAGGCTNGGTAAAA  
TTTTAATTTTTCTTTGGGGGANCCTGGGCGGNACCACNTTAAGGNCNAATTCCACCCCNNTTGGNN  
GGCCTTTTAGNGAACCNAACTCGGGACCAAAATTTGGNGAAAAAAGGCAANANTNTTNTCTTG  
NGAAATTGNTNCCNTNAAAAATCCNAAAAAANAACCGGAACCTAAATGTAAACCNNGGGGCC  
CATNGNNGGCCNCCNNTTNTTGGNNGCC

SEQ ID NO: 421 ACTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGTTTTTTTTTTTTTTTTTTTTTTT  
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGGGGGNTTTTAAANNAATTNTTAGGGGGGAAAA  
AAGGGNNTTANTNTNTTGTCCANGCTGGCCTTNAACNCCTGGNCNNAACAATCCTCCTGGGTNG  
GCCNCCCANANGGGNNGGGTTAANAANAAAAATTTNTTNTNNNTTTNTGGNAATTTAAANTCA  
AAAANCAAGGGGNNGGGGANTTTGTTTNTTTNAAACCNNTTTTTTGNNTGCAAAAAACNGCCN  
TCTNACTGGGGCCNCCCTGGCTTGGCNGTGCNNACATTGNTGGGATTTNTTATGCCAAAGGAN  
CCATTTTGGATTAAAAAACACCTCCTTTGAAATTAANNCCCTTTTNAAGGGCTTATTTTCNAANA

GNCCATTTAAAAAATNGGGGGTTGGACCCNCCAAAGAAATTGGGGGAACACAGTNACCCANC  
CAAGGNCCNTTTTAGGGTTTGGTTAAAAACTTGGCCAAACCTCCCCTANGGNAAAAATTTTTTTA  
GNAANTNAAATTTTTTAAAGGAAACCNCCCCATTTGAAAAAATAAATTTTCCTTGNGAA

SEQ ID NO: 422 ACANTATGTNTAATNNTTAAATGTTTTATTATTGGAAAAATAANGCGTGTA  
TANNATGCCAGGGACTGNCAAANGACTTGATACAGGATGGNTANNCTTGTCAGCTAAGGNCACAT  
TGNGCCNTTNGACCTTATCTTCCTGGACTATTGAAANCNAGCTNANTGNATNNAGNGATATTACT  
ATANCGATTGANNNGGCANTAGTTAAAGTNATNAGCATGATNAGAGTTNCTGNCAATCATGTATT  
ANAACCTGATTNTNAGNNTNACANAAATGTCAGANTTGCAGCTATTGCNGGAATCCAAAGTNATG  
GCCNGCTAGCTAGGTTAAAGATTGGTTTAAATCTGGGATNTTGCTTTTNCCTGTNGACATTGCTT  
GANGACATTATCTGANAGACAAGTTTGTNGCNCAGTTGCCTGNTANAAAAACCTTTGNANACT  
TNNTTNTNCAANAGCCATTGGNAAAAATCCGAGGAGNTTGGANNGAAAAANCNNNGCTTTAAA  
ACTNGTTGCCATTTTAAANANNANTTAATGTATCGTNAANTTTTAAGCCTCNANTTTACAANTCCC  
GCTTNGATAANAGACCTATCAANTNTTTTGCTCAAANGCTTNTTTACNCTATTA

SEQ ID NO: 423 GGTACTGTNGNNTCATCHNTTGGGANNGCCACACCNACANNCCNCCNGCTTTC  
TATGCNCNACATNANTGCCNATCNNTATGGNNANCNNCAGANCNNNGCTCNGCNGCNGCNCNG  
NTTNANNNCNCCNCCANGANGATCNCACANTATGGACNTGCTCCTGTCTTTTAACTGTGGCCAN  
CGTCNNGCATCACGATTTGGAGTTGCTGGCCAAGGTGGCTCTGATAANCAGCCNTGGTGTNTANA  
NGATATTTACGAAGACTGGCNTTANNGGACCATACCTGNANTNTTCTANCTACNGNAANCCCA  
TTTTATNCCATGGANNNTNTTAAATCAANGNTGCTNTGGTCTGAAGCCCTATATGCTGGAGATGG  
ACNCCTNAATNAATNATTAAGGGAANNCCCTATNCCTGANGTGGGGTGCCNTTTCANAGACTT  
TACNTAANTATAGACGGGCTAACCTGCAAACCATNNTGAGAAATGACTCTTNTCNCNTTNGCAN  
GGTTTTCCAAGATGTCNNTACCANACNCNTTTCNTTGAGAAGGNTTNTTCCCCCTTAATANNNCC  
TGTTATCTTCCCTTTTCNCTGAAGGNAGATCTGCNTNAGGGTCCNTAAAAAG

SEQ ID NO: 424 GGTACTTTTTTTTTTTTTNTNTTTTTTCTTTGGGAGACTAAAAANTTTATTGCAT  
CTTTAAAGCCTTAGGCCGTATGACTAAATGANTAGACTGNANTGACNGCGGGGAGGAAGAANCA  
NANGAAAGATNTTAATGAGNGGTCNGGTTGGGGGAAATAANNCGAANATTNCNTNCCAGGGTG  
AGTCCTCACACTGGCCTNATGCCCTTGNTGANTTGNCCNCCAAACACAGGGCTNGNTNACTTNCNTT  
CTGCACTAGCAGAGAACTTGCNANATTAGGNNNACCTNACATNNCNGTGTAANTCCTTTCCC  
GCANAAATGNTGTNGCTTNTTGNAGGNTNAGATAAGTGNTCCGGGGCTTNAGATNGGNGGANGG  
ATTTCTTTCCCATNGCTAAAGNGAATNCATCCCATACANNNTGNTATTTANCNTNCCTAACTTTTATC  
NNTTGGGCTNNGGTGNCCTNTGTGGGGTATTGGNACCCAATTACATACNNAGTTNGNGTAAATTT  
ANCCAATGNNACTNTNATCNGGGGAGNAACANAAAGGATACCTTCATNTCNCAANNTCACNTNN  
NNCCTGGTTTCCAACAGAGGCCTTTTCNTGNTGGNNAAATNANTGTTTTACNCNNTATNTNCGGNG  
ATCACATNCTGGGNNNTNCTNTGNGNNTCTAACNCCNTT

SEQ ID NO: 425 TCNAGCGGCCGNNCGGGCGGGNTCTTTTTATGAGAAGCGTATGGCCACANAA  
GNTGCTGCTGACGCTATGGGTGAANAATGGAAGGGTTATGNGGTCCGAATCAGNGGTGGGAACCA  
CAAACATGGGTCCCCATGATGCAANGGTNTCTTGACCCATGGNGCGTGTCCGTCTGNTTCCGNGC  
AAGGGGCTTTCNTGTTNAGACCACGCGANAACGGNTNAANAGANTNNATAATATCTCCCTAGGT  
TNCACTTGANAGNTNCATATNAGATCGTNCCTATNNTTNTNNGNTAAAAATGGTNAAGTNG  
ATTNTCCTGNACTGACTGATACGGANNTGCGTTTGCCNGTCTGGGCCCCNACNAGCTTGCA  
ATCCNNNGGATTTTCAATTCNTCTTTAAANAATATNANTGTCNGCCNNTANGNTTGNACNAATNC  
CCTTACANTTAAGATNGGGNATTAATTCNTGACCNATGCNNTCNCATCATTTNGANGTTTTATTC  
ANTTNNNNNNCTNGAACTCATACCNGCTGCGTNATTTTCGNNGANANANTGCNNGAATTTNGTNN  
CNNCACNACATATGGCANNANTNCTTCCCANCTGGTGNCCNACNTAATTGGAANCNTACCTCGGAA  
NCNTNNTNGGTNAAAGTNNAGTNGTANNTNCTTCNTGGG

SEQ ID NO: 426 CACTTCAGCAGCGNGGCGGGAACTGGGGGTATTGAANAACNGGCAACCNC  
AAANTANNATTTATNCAGGGGNCAGAACGNCANGCTAGACTTNNCTTCCACTGGTCACAGNATAA  
NGGCAACCTGCAAGACTNAAAAGNAGGAGANGANGNNGAATNGGAAGNATCANCGGACCGACG  
GACTGGGAAACCCANGGCNGNNAACGGGNGAAGANNAACAAGAAATCCNCCCTGAAGAACC  
GAANGANGGGGACAAGAAGGGGNNACCTGACNCACACCCCGGAAGANGNCGAAGGGGANGCNC  
AGCNGGAAACCGGAGGANNANANAACTGGGGAACNGANAACAANAAACATACGGGGGCCGG  
NAANGNGCCCCGCAACANNATGCGGGGGNAAAAGAAACCACGCCCCGNGGGGNGGAGNAG

SEQ ID NO: 427 ACCGATGATACTGNCNCTTGNCNTGANTATNTAAACACTNCACAGTGTTNAT  
ATNGGGAANATATNGGGAAGGAAATATNTNNNTNAAANATGAACGCTGNCNTNATGTTTTNTCT  
NTTTAAACTGGCTCACTTANANTCTTTNAGGATGGGANGNNTCTTNTTGCTGTANAGCGCTCN  
NCCTTTTCATCTGATGTTTCNGCGGTNNCANGGACCTTAGCTACCTATCANACTTTTCCAANAGTAT



GTN

SEQ ID NO: 433 CATATATACCCAAGNGTGCATATCTTGATCTGTATGCTCTTANATGCGCTTTA  
TNTACAGCTACNGCACATANNGCNCACATANCATNTTTACNCACAANGGTTGAAAACCTGTGCAT  
GATTNNATATCATCANCNAGCACGNNTGCTGG

SEQ ID NO: 434 GGCCGAAGTACGCGGGCCTGAACCCAAGAGACAGAGGTTGCNGTGAGCCGA  
GATCGCACCATTGCNCTCCANCTGGGCANCNAGCANAAAACTNTGNCTCAAAGAAAANNAAA  
NANTAGANCNAGACGANAATGGCTTNCNGGACAGGAGCATTGCTCATTTGCTGGCGGGACNGTTC  
NANAATCANCCNTGGTNTGGTCCTTCTNCTTACCCTGGCTNGTTTTNTNCAANCCACGTANNTTN  
NTAANACTTTNTTGGAAGCNAAATAA

SEQ ID NO: 435 ACTTTTTTTTTTTTTTTTTTTTTTTTTNTGTNTTATTTTTTTTTTTTTTTTTTTTTN  
TTTTTAANGCTTTGTTGTTGGATNTATGGANGATGGGGATTATTGCTAGGATNAGGATGGATATT  
AANAGGGCAAGGACCCCTNNTAACTTGTTAGGGACGGATCGGAAAATTTNTNTTGGNAANNAGN  
AAATATCATTCCGNNTTGATNNNGGAANGGGTNTTAAACGGGTTNGGCTAAGGGTTAAAATTGG  
TNTGGGNCCCCCNAGAAGNCNNGGNANAANATCGNTAATGNCATTANAGGNGAAAATNAACAN  
AANTNANCCNNNAGGCNNNTTTNATNTTGAANANAANGGNGGANAGGGNNTTTATNNGGAATG  
GGAGGTGNTTCTAGGGGGTTNTNTNAAACCNTTTCGTNCAAAAAAAGGAGGNGGAANTGCTTTN  
TAGGGCTTNANTANTNAANGGCACATNAAATTGAAAGGTAAAAAANNNTGNGAGGGGGGGANT  
GTCTNCTGANTAACCTNCTNAAAANTAGTTAAANNAANGNTGACCCAANTTTTGGGATNGCAGA  
NAAATTAATAATNCTAATTACTNGGGCCCCANANAAATNANATTTTNGCCTNNCGGCGAGNANAA  
NNCCTTTA

SEQ ID NO: 436 ACCTTGATACACATAATCAGCCTTTTCAAAAATGCCTGACAAGAATTAGTCTT  
TCCTTTGTGCTAAAAGTCTTCCCACCCATGGATGGAAACAGGCTGACTCCTGGAGGGTCAAGCAA  
GGGGTGGGGAAAGGGGAACACANTNCTTTTGGGAAGGCNAAAGCAAAAAAGGGGTNTTTTGNC  
CAAACCAACNTTGGGGCCAGCTCAAANGGGGNCCNAAAGCNTNCCCCNAAAAAAGGNTTTCNTTTT  
TTTTTTNTAGGGCCCTGCACTTTANAAATTTGAANGNTTGNANTTAANANTNNGGACCCNCTNANA  
CTTTTTTCCANANGGAAAANTTTCGGGGGCCANTCNCCCCAAGGNAANNNGGCCCTTTCNNAAN  
CCCNNTTTTGNAAGNAAGGGCCCCGGGGGTCTTNNCAAAAAANNTCNNGGCNNCTGGGGGG  
AAAAAAAANCCNTGGGGNAAACCGNCCNCTTTANAANTTGGGCCCCCTTTGTGGGGGGCCCCN  
TTTTATTATGGGGGGGGGNNTACT

SEQ ID NO: 437 CNGACGGGCCCNACAGGAGCANNACAGGAACACTACTGGNTNCTGNNAACCTGN  
GGNTGCTNATGTGNCCACTGGGCACCTTATTGCNAAACCTGAACCAAANANACCTCCTTGNTTCAGC  
TTGGGCCTGCTGTCCAGCTTCCGAGGTGCAGCAGGGTTGTGGGAACAAGAGACGACTTTNAGGAT  
NAAANGACCAAAGGANAAAGCTGCCTTACATGATTTGATTGGGGCCTAGGAGANTGGAANTCACNC  
TTTATNTTNTAGAGAGNTNTTTNACTAATGNNGNAGGCTGAGGNGCANNCCTTNGAATATGCCTT  
ANANGGCCGNACGCGGTGGNTCCCCCTGCAATCCCNNTACTNNGTGAGGCNAAAGGTGGGCNNGC  
CANCTNNGGCTCGAANTCAAAGANCCNTCCNTANANTANCNNNGGGTGAANCNNNTTCCNTAN  
TNNNNATCCAATNATTTATCNTNNCATCTNNANCTCATNNNNCTNNTAANCNTNANNCTNGTGNA  
GCATGTGGAAANATGAAATT

SEQ ID NO: 438 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGATTTAATTNTNAAN  
CAAAAACANCGGAAAANGGGATTAATNATNNGGTTGTTANACNGGGNCANAAATAAACNCAAAA  
NNAGCCCAAAATGTCNCAGGACCGGGGCAGAGGACCAACATGGGCNTTTTGTNNATNACCANGG  
GGGGACCNANAGGGGANCGGCNATNAAAGGGCAATNAANNNTAAATNCATTGAAACAAGCTTT  
ANACAGTNTCNTGCANTCCACATNCTTGTACCTNGGTCGNAAACCCACNCTAAGGGCAAATTCAN  
GGCNNANTGGCGGCCNTTNTAGNGGATCCAAACTGGNTCAAANCTGGGCNAAATCATGGNNAT  
ANCTGNTTCTTGGGTGAAATNGGTATCCANTCACAATTCACCAACATACCACACCGGANNCAT  
AACGGGTAAANCTGGGGTNCNNAANGAGTGNNCTAAGNGAGATTAATTGCGNTGGNTAANTGG  
CCGCTTTC

SEQ ID NO: 439 ACTATGTCGATTTCGACAGAACANTTTTTANGATTCTCGGCCTTGCCCCCTTCAC  
GAGCCGCCACCAAGCAGGCGGATTCTTGGCCACCACCTCCTTCTGGGAAGTCTCTTGAAC  
TCAAGACCTCTTTATTTNCTATCATTCTTTNCTAGACACACACACAATCAAGACCTGGCAACTGNT  
TTTGAACAAGAGCCATNAGGTANCCNTTANTACTTGGGGCCNCTTTCTNAGTTNTGAANTATTTCC  
AAANCTTTTGGGTATNNAATAANAGTNAAAAAGGCAANCCCGCAAACANNGNANGTGACTTTG  
GNCCTTTAAGATCTTTNNNAAATNAGTGGATTGNATAGTAANNTCAACGAAANGGATATNGGAAN  
GAAAAACAANANNTCCCTCNTGGGNTCATTACNTAAANGTTTTTACNTGGGGNANCNATCTA  
AAAGNGNCTNTGTANGCCCTGCAAGTTGGCTGGGNTTGANCAATTTTNGAGATTTANAAAAAAA



ANTTANGTTTGNATAACC

SEQ ID NO: 440 AGCCGTGGTCCGGCCGAGGCACCTTTTTNTTTTTTTTTTTTTTTTTTTTCGTTTNCN  
GAAATGTTCTGAAGTTAACTCATTTTTATTCTAGGATTCGGATTNCAACATTTNAATTTCTTNGGAAT  
ATAAGNCACCTTTTTGCAANCTAAAAANTAGAATCAAACCTANNGGTGANCTAGTCTCTAGGCAT  
CCAAGGCNGATCCTTGGAATCNTGANCANAAANGATNGACATCCTACANGGTGCTNGCAATACGGC  
TATAAACCTTCTAAANACTTNACNNCTTACATTGTTNNATTAGGAAACAGCCAANAGGCCNGCCG  
NTAAAGGAATNAATNTGCNAAACATNTTGNNTTANNNNGGGACAAAATGNTGNCAATTATTTGNTG  
GCCACCTACTNGTGAACCTGNNANCANAAATGGTCCCAGGGACTGTNAANCNNAANGAAGGNTTT  
TACATTTTTTGAAACTTGCTCTGCAGGATAAAAAAANNCCNGGNTGCTGTCNGTNTAAANNCCNG  
TGANANAANTTTNCCGGTTGANCCCTCNCTGGACTNTTGNNTTGNCNANGGGCCNCCNCTGGACN  
TGNAGGGGGGNGGGGATATNTGTACCTNCCCGGCCGGCGTTTAAAANGGGNAATTTCANCCNACTG  
NGGTCNNNACTTNGGGGNCCAACCCNGNNTCCANC

SEQ ID NO: 441 ACACGGGGGTCGGTGGCCGGCAGTCATNTCNCGGCCGTNTCANAATTATAAG  
GCTGNCTGCAGAGATTNCGAANAAATGGCAACANATGAAAGCGTCANCANCTTTAGTTCAGCATC  
CTTGGCTGGGGAATATGTAGATTTCACTTTTACNTTGANAATTCNTCCTNANNGAAACAATTTACT  
ANATNGCTTGAAACTATATGNNTGNATAAATNATNCNAAAGTTCCANTATTGGCAATCATTGGG  
GGATCCCCTTANTNGTTCAANGAAGCCCTATTATTNCTTAATNCTGNNTNACCTGNGNTTTTTAAT  
TTNTCANTNATNCCTTANTANNAAAAANTTCNANAATTCAA

SEQ ID NO: 442 CGTGGTCCGGNCNAGGNACNAGAATGNTTCATGAAATCCGNTTTTAAAATGA  
ACNTNTNTGNGNGCCACANTTCCTANGACTGGGGCNAGGNCNNTGNCAAGTNTTGNNTTGAGG  
NTNAATCTNTNAANAAACNNAATTCCTGCCNNAATGCNCNCNNAACATNAGTNANCCNAGTGNGT  
CNTGCCANGGATNCTTTGACTTTGGTTTGCTGCTGNTGCTNNGGATATTGGGAGGGNTATNCTTTN  
CANGTTNNAAGAAANGGNTGTGGGTTAANGGCTGTCNTAAAAGANCCCTGGCTGTNACNCCANCT  
GANTCCNGATTGCGTTTGTACCCNTTTGNAACTGACCCGNTAATTTNAACNNTTTTTCNANCTTTT  
TTNAAGCNTNTTTANGAAGCCTTCCCGGGANGNAATTTTTTCCAGGTNATNTNCTTNNCCGGGCN  
GGGCCGTTNAAAAGGGGGAATTCNACCCACTGGNGGGCNTTCTAATGGGANTCC

SEQ ID NO: 443 ACTTTTTTTNTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTNTTANGGCCAN  
NTTCANNCCCTGAANTGNCANANNNAANTNAAAATNTNAAAANATGCACCNTGAAGGCTTTGAG  
NCGAAGAANTNAATCNTNNCNGGGGATCTNGANCTNCAANGACTGTGACGGNAAAGGATGNTN  
TTTTTTTTACCCTGGCTTNNCNCANNTATNNCAAAAACAAAATGGGAANCAACCCNCAATTTTTT  
GGNAANNAGAAAAACACNACTAGGGNACCCCAAAAAAACCCCATTTATTTTCTTTGGAAAAAG  
GNNGGGGNCCANGNTAAAAANTGNNANGGGGNNAAATTACNTTNANNCAAAAGATGGAAACNT  
NCCTAAAAGNTTNCNCNNTNGGANGAAAGNGGGTTAAAANNCTNAAAATCTTTNTTNGGGAAAN  
AGGGACTCTAATNGGTCCNTTNGGGCTNCTTNAANGGGGAAANGGAANNAAANANTTTNCTTNA  
AAAGGGNGGCNTTNAAAACCCNNAANNGGAAANAAAAAANTTTCCNGTTGAGGTNNGNNAC  
ANNNTTTTGAAACCNGCCAAATTTCCNTNTNGGGNAAATNTTCTTTAAAAGCNANAAANANATT  
TTTAAAAAAAATGGNTTNGGGGGGNGNCC

SEQ ID NO: 444 ACCCGGNGCCCNACGGNGCCNNACAGATGGCTGGNTNNGACATNGGGCNA  
NNCTGCCAGNTGGAGCATTGNCGGCNCNCCGAGATTTNNTTNCATNTGGGGTGTGATGATTGTTCAT  
GAATATNTTTGCCNGAACNCAGAANCTGGNATTCTCATGGNTGAGNTGAGGTGACTGNATGTCA  
NTGAGAGACTGAACACANATCANCATACATCTTACCCATGCTCTTTCAAAGACTGTGCTAAGAGA  
GAACCTGTGCNCATTCCCTTCCNTATNGGCAC

SEQ ID NO: 445 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTTTTTTTTTTTTTNTAGGNTTT  
TNTTTTTTTTTTTTTTNCAANCCTTTCANATAANTNNATTTTNTTCCCTTCNACCANAATAATGGNAT  
ATCCAAATTTTGTCTTAACCGGGGGGTNNGNAAACACAGGCTTNTCCATCCCACTTACAGGCAAG  
GNAAGGCTNNTTNTTGAANGGCCCAACCATGGACCCTTNGTAATCCAANCNAAGCNACCCCC  
TNGCCGGGNTTAATNTTAANTANACATGGGCGGCCAGTTAATNAAATCTNNACCCAAATNTTAAA  
NTTTNCAAGNGGNTTCCNANTATTTGGCATGNTNTTNAANAANAAGGGGANAGACCCTTTA  
ANTGNANTTNCNCCCTTGGGGNCCNTNAACCNNTTTGGACAAANCCTGNAAATNACAAAAGG

SEQ ID NO: 446 ACTNCTTTTNTTTTTTTTTTNTNNTTTTTGNTCAAGATTAATNNCTATNGACAAGG  
TINATAAGAACANATATTTAAANTCGAAGGCCANTTNTTAGGTCTCAITTAGCTGCTTATTTNGTT  
CACTTGATTTTACCTTNCCTAGNGNGGTGAGNAACTATCAAGAAACAAACCTGTGAAAAACCT  
GNTAACATTCAACANATATTTGGTATATATANCGGTCTNGGANGCAAANATATTTNTCAACACTTA  
ANTGGNGNANCAAAGNGNGTTCATNGGGANATAAACAGGATNGNTTTAANNTTGAGANTTTAATA



NA

SEQ ID NO: 447 GTANTCTTCNTNTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTT  
GCAGAAATTCGCCCTTNCGAGCGGCCGCCGGGCAGGTACNCTTAGAGGTANNTGTNNTNANANAN  
AGNNNTGTNNNANTTAAAGANNCTAANGTCNTNATNTNTTTNATAGNCAANNTACA

SEQ ID NO: 448 GGACNTTTTTTTTTTTTTTTTTTTTTTTGGAGTTTTATTTANTATNGGGAGCAAAT  
TGNNTTNTAAAATGNATATTGNGAATAAAACNGCNTTTNNNCNTTTAANGGTNTAGGGCTGTNAA  
NNGNNNCANGGGGNGNNGCCGANNNAAACCATGANCTGGGCTGGGTTTNTATNTTTGATGAAAAAN  
AGCCTNNNCGCCTTTTGATTNGGGANANAAAAAANNNGNCATNGACCTTGACTATNNCTTTATCTC  
CNGNCACCTTTNNAAACANTAAATTGCNNGNCNNGGGGGGAGGGCTNTTNNCGGAACTAAACTGT  
GAGCCCTTCCATTTTTTANGAGGGGGANGTGANAAAAAGATTACACNGTGGNTTGAGTGNCCCTNA  
GGAANAATTNGGACCTAACNTGCCCGGAAANGAAGNNATAGGTCTGATNCGTTNGTANAAANACG  
ACAANNNTNCACCTCCCCCTCCTGGTGATGCACTGANGGGGACAGGTGGAGNGANANTACGNNTA  
TTGCNACCAAGTNTCTGTCACTNCTNGGGANGACCGNGTNTAAAGAATTTTGAAATAAGCTCTTTA  
CNTTGCCTTTTNA

SEQ ID NO: 449 GTACTANACACATCNGGGACAACCNCCATTTCCGGANATGATGCCGAAAAACNC  
AAGGCCANAAGCNNAAGCNAGGGGATGGANAGTTTGNNGGAAGNTATTTCTTTACCCANNAATGA  
CCTGNTGCAAAGACTTGATGCNCTGGTANCTGANGAACATCNTCACNGTGGACGCCANGGTCTAT  
NNCTACGCTNTAGCGCTGAAACATGCNTAAGCAAAGCCATTTGANNGTGCCCTTCTTGANATTTTA  
NNCCNAAATATGACACTTGGCCATNTATTGTNAACGAAAAGNNCCANGCCTTGTTTNTAAGACNT  
TG

SEQ ID NO: 450 CGCGGCGAGGTACTT  
TTTTTTTTTTTTTTGGACNCCCAAACCATCCTTTATNGGANNATNANTTCANGGNANCNGCANNA  
AAAAACATTTAGGNGGAATTNANAATTNCCGNTNAAAAAAGTNGCCCNCCAAACANAACCAATT  
TANNAAGNCAATTCATNAAANGGNATAAAACCNNTTGNNGGGCATGANGGCANGGGACAAAGC  
TNAACTTGCCCTGGNCCTTTNGAANCCNNGGNAGGNGANCNTTTNACCCAAAGAACCCNGAA  
NCCCCCGGGGCAAAAAAGAAAATCCNCTAAAAAATTTTAGCNAGGGGGGC

SEQ ID NO: 451 ACGTGCCGCAGGAAATACTCCGGTAGCAATCGCATNATCGGTGCCAAGGACC  
ACCCATCCTTCCAAATGANCGNGGCCNAGGTTGACATAGGTAAACNGGNAGNNTAAANGGCCAATT  
NANANNTGNNNCNTTANGGGGGGCCNTNNCANGNATGGGNGAGNAAAATNTNCCNTTNTCCNA  
TNGGCCAAGGCCNANGGCATCNTNTNAAANAACCTTTNGACNGGAAAGAATCACAAATGTGGAAT  
NTTGNNTAANTAANTAATGAAACCCAAAAAATTTAGCNAGGGGGGC

SEQ ID NO: 452 CNCCTTAACCGTGGCTTTNCCGACGTACANTCCAATCGTCTTCGNGGGGNTTN  
CNCTTAGCCGANGAGTTCNCNACNNNTTCCACAAATTTTAAAGANGAGGTAGACCCACCTCCAT  
CTTAAANACTTT

SEQ ID NO: 453 AACTACCATCTTTACATCAAATNGGGGANCCTGGAGGTAGTGAAAAAGCTA  
TTNGGATTNAAAGTGNNCAGAATTCNTGTANACCAACAGCAACTGANCCNCTGTGTAATTAANGGC  
NNTCCANGGGNAATCANCTANTTCGAGCNTNTTNCGCTATTTAGGCT

SEQ ID NO: 454 GACGGAGCAATCGANGAGGCATAACCACACTTGGGGGTGGGCTATAGGGGC  
TGGGAAAAACCTGAAAAATGAACTGGCTTTTCACTGGAGGGCCANGGGTTTGGAATATTTGGC  
CAGNCTTNGAAANGTNTTTAAAGNCAAANTTCCCTTAAAGTGANTCTTTTCTCTNAAAAACANCA  
NTCANCTTCCATTGTNCCTTACANAAGGNCCCCCTGCGTTCTTGCTTGNATTGCTTTTGCGCACTN  
CCTTGATGATNANAANGGCCCAATTATTGATNGCCCNCCNNAANCCNNANTCNGGGCCANGG  
CACNNACCAANGTTCTCTGTAATCTNCTGGGANAAAAGGCTTTGGNACNNTAAAATCNCCAATGCNT  
TNAGAAANAANGNTNATTNCAAGCCCCATGCTCCACCCTGCANTCGTAAACNTCCTCAATTNANG  
GCNGGAAGGGGAAAATATTNGGAACCTCGGGAAAAGGGGGNTTCTTGGGCAAGGAAAANTGTC  
TCTGCACTCCTTNGNNGAANGGTTTTNAATTTTAAACCNCTGACNTTNACCNTGGTNGGGNGGA  
ATTTTCTANAAACANATNAGGANANTTGGNTCTNTTCAANTTACAAGGGTGTNTNTTGAGCA  
CAGTTTCTNTNNGAAGNAANCTNCC

SEQ ID NO: 455 ANCNTGGCCGCGGNCGAGGTACTNNCTNNACTGTGAACGGGCTCCAAAGGA  
CATGGNTCTGCANTCAAATAATNATNAAANGGACAGGCNTNGCNAAAATGCATNGGGNCCNAC  
TAANCNTNNNCACNATCAAGGNACCAACACTNNANNNGNTTTNGGCCATTTANGTTGCNAGGT  
ATNTGNGGNCCTGCTNCNTNTTTTTTATTCANNGATCAAACNNTCNNAGACNTNATNGCCTATCANA

TGTGGTAATCCNNAACNCTGCTGTCTCNNTAAANTTTTCATANCATGACTTACNGTTNGANNAAACC  
CNGNANTTGTAGANGTCCATAACTGGTNGGAGAACGNCGGTANACATNTCTTTTTTATAGAGGT  
CGCTACGNTCTGATTTNCGCNCAATTNCNATNCTNNTCNANAANTNNCTTTNANCTTNATTNNAA  
TGNTNGACCATACTGGGGGNCGGANACCTTAGGGANATACATCTTTTGTGTCANAAAACANAT  
TAATTTAAATCCTTGGCCTGGGGGGGCTCTCACACINGTGAGNTTTNNATTGNAATATNCNGCTNG  
NCNATNACGGAGGNTGCGNNNNATCTTGTTGCTNNGCNCTCTGCTCTTANTGTANANATNNAG  
GNAGGTCCG

SEQ ID NO: 456 ACTTTTTTTTTTTTTTNTNATTTTTGNGGGTNTTNGCATCCACAACNGTCCAAN  
ATNTTTAAACTTNTCTCCNGTAAACNAGATGAANAGGCGNCNATCCTTATNAACAAATTGGNNC  
CATNGGNCANTTCGGCAGGGCCANTGATNGNAATGCCTNTCANTTTTTNTTGAANACNCACAGTTT  
NTTCAAATNTCTTNAATGAATAAATCTCATCCNCCCTTCTTTCCCTTTTNAATATNTCACNNCCA  
TTTTNNTTNTTAAACCTTCTTTTNGCTGTTTCNGNAATTCCTTTTTTTGGAAAGGAAGGCCANTG  
GCCCCTTAANAGAANGGGGGCCATTTTCTTNGACCCCNATTTCAINTTTCTTGTCTNAGTTNAA  
AATGGNNTNANAAATTTNCCNGGGNANATGNANCAATTTCTTGANTCTNTGAAAANTTAAATTT  
CCAANCTNTCTTTTTTCTTAATNTNNTTNTTGTGTCNNNNNNNACTAAANCAAAGTGCTNNTTCNN  
AAATTTTATCTNANNGNGTNTATGGTGGTAANTCATCAGATCTTCTGNANNGNGNNNTCTNTANA  
TNG

SEQ ID NO: 457 CGTGGTCGCGGCTCGAGGTACACTGGGAACTCCAAGAAAAAGCTTTGAANAG  
AAAATGGAGGAAGCNCGAANCCAGAACTCTAAATCTTGAGACAAGAAGACTNTCTATNATGAA  
NGAGATTCTTTNTGANCAANACCNAACNCACAATGTGTGGGAGAAATCTNATGANCTAATACCAT  
GCCTGCGAAANNAAAANTGCNAANNNAATNTTCTNTCCGNGCGGCCGCTNGAAA

SEQ ID NO: 458 TTNGTACTGATTNAAAACTAATCACTTAAATGTGCCACNCCGCAAAAAGAG  
AAAACCAAGAGTGGTCCACAAACACATGCTCCTTTCTCTCNGAAGGTTTTACNANNCATNGTAAT  
ACATAACCCANTCTTTTANTANTAACTNAANNNGCCANTNGAAACAAACAGTTTTGAGACCGTT  
CTTCCCAACCTNNTNAAAAATNGGGGGGCGAGGTATNNGGGGATAATATCENNTTATATCCGACNT  
GANCTTNTCTGGGCGAAACTTGGGTGACCTTGNNANNNTCCAGCTTNTCTGCTTGTCCACTNNCTTT  
TGATGACTTGCAACAGGNATTNGCTGCTTCTTTGNACGNACNNGTNNNTTCCGCCCTAGGCGTGG  
NTAGNNTTANAAAGCTTTAACTTNGAATGGGGCTTTNTNANGGAACCCATNATNAACNATTGGG  
CGGTTTTNACGATACTNTAANATANTTAAAGGGCNCCTTTNCANNTTGTNTTNCANANTNNGT  
NAAATANTTTTTTTCTATNATNNTTTNTAAATCTTAACTTAAATNCTNCNCCNTGTGTNAACTCCT  
NTANNATTGGGNATNNTCNGACCTNTNTATGNNT

SEQ ID NO: 459 GGTACGCGGGGACTGCAAGGCGGCTGCGANAGAGAGGTTGTGGCGCTANTTTC  
TCTAAGCCATCCANNGCCATCCTCGTCGCTGCNNNGACACACCGCTCTCGCCGCCGCTTGANTGA  
CCAATATNACCNTTCGTGGCACCCCTAAAGGGCCACAACGGTTGGGTAAACCCAAATCNNTANNNCN  
CGGATGTTNCGNANATGATCCTTTCCGCNTTTCGANATANAACCNTCACANTGNTGGAAACNNN  
CCAGGGTNGAAACCANCTATGGAATATCCANNNNGTCCCTTTCGCGGGGTAACCTCCNACTTGTGNG  
AGNNATTNTGGATTATTNCCCTNATATGGCNCANATTGCCCATCTAAGGGATCCTGGNNATGNAA  
CCNTTGGCGCTTTGNGTATTTTACNAACTTNGATCCCCCTCNNAANGGGATTGTGGGCCATACCT  
CATGAATTGCTTGAAAGNTGNCNTTTNNNTANANTACNNNGNNANAATTTGCTNTCTGGAATCTT  
AATATAAAACCATTAAGNTTNGGNGANTACCTGGGNGATGTGAAAAANACANTGCTNNGGAANT  
ANACCNATNTAAANGNGTTTTTT

SEQ ID NO: 460 ACCGCGGGGCTGNCTCTCTTTTNGACTCAGCCCGCCTGCACCCANGTGAAA  
TAAACAGCCATGTGGNTCACACAAAGCCGTGTTNGGNGGTTNTTTAACATGNACCCANATNAAAT  
TNGGNGCCATNANTNGGATCGGGGGACCTCCNTNGGGAAANNAATCCTCCGNCCTCNNGNTCTTT  
GCTCCNTNAAAAAATCCNCCTACAACNTCAGGTCCTCAAACCGACAGAGCCCANNAACATTTA  
ACACANTTNAAAAATNTGGTAAGCAGCCTTTTTTATTTTTTTTTTCAAACCTTCTTCAATATCCNTA  
AAACCATTTTTCTCTCTTTNANTCTNTNCCNNTTNNACTTAAANTTTNTCNTCTCTTNTTCTTAATTAT  
ACATATACCNTATTNATTTTTCTNTGGTAAGANAAAAAAGAANANANNNTNTTTATCCCGTNGG  
GCCCCAAAANTTTTCGGNCNNTTTGGTCNAAAGNNTNGNAAANNGNAANTCTTCCCTNNGGNTNNC  
TTAAATAAATGGAAGGGANTGCTCTTNTNATTTATACACNCTATGATTATANGGGGTGNTANTAA  
CANTCGGGGAAAANNNGTTTTGGATTTT

SEQ ID NO: 461 GGNACCCCCGAGTCCNTGNCTGGCATACTGAGAACNACCAAACACACACCCA  
AGCTCGGTCTCCTNTTGGTGATTCTGGGGAGCANATCTTNATNAAGGGCNCNCCGTCCCATGAGAG  
GGGGAAAACATCCCTCTCTGNAATNTAAGGAAANTTTTCNCNTATNANTCTGTAAAGNAGAA  
TTTCTGNATTTTCAGACTANTGNNAANNNTNANCCCCCNAGAGANNNTNATNTCTTTNGGGCCT  
NNGTTCTTACCTTGANANTTATATAAANGAAAGTGTGATATATTNATTGCANNAGNGTTCAANA

AANAAACATCTTATGTTATGTGNANNTNAANTANNCTNANNTGAGGNTTNTTAATNGANNGT  
GNAATGNTATTNAATCACTAGGTTNTTTNTNATGCANTAAANCCTTNTATTTNATNCTTTNATTTNC  
TNTTGNNAATGTGAGATANTANNGTTTTTTTTTTNGCNTNCTGGNATNTTGGNTATNANTATAAC  
NTATNTNCNTATNNTGTANTATTGGATTTTNNANNGGAGAATNTTCNTAANTCCTTTTNTCTNNTTATT  
TNTNTATTGATTTTATNCNCAACANTCTTNTCTT

SEQ ID NO: 462 CGAGCGGCCGNCGCCGGGCANGNACTTGAGGCATTACCTCCTCTGAGNCACTGA  
AGTCATAGGGGNCGTATGAANCCCCATNTTTGGCATNNGGNTGGCGANACNTACTNCTCTTGCTA  
TTNNGCNCNNGGTCNTCTTGATTNTNTGCTCTTCATCTNCTGNCATGAATNTTCTANCTAAAN  
TNANTACGGTTANTGTNCNTCACNNTAACCTTCNAAAAATAANCAANNNTNTGNCNGCATCCGNAA  
NNACNCTCTTCTATTCTGNAGNGACACAGCTANTGCTCATACGNNGGNTTCTGATGGGCTGTGGCN  
ATGNNAATCAAANGATTTNTAGAGTNCGGCTTGTAACCTGGNGAAGANNTGGTCGNCGGNACN  
NNCCANANNCNTATTCCTTGCCNGNGCATGTCTNNAAANCATTTTNNAAANGGTANGGGGGGNTC  
GCCTTCCGTTGANATGAAACGCTANAATATTNAAANCCCTTTGGTNACNGCGGGNGCTNNTGTAT  
GGAGCCCTCNTTTTNTNTAAANTTNGGGGTAAAATAACCCTTTTGCNNGGGCTTANCCGGCNTGTG  
NTNTATTATATTACACCGAAGGGGCCCNCTNTNTTTGGGGCANGGCCTTTTNTTTTAAANGC  
NTTNCCTGTCTCCNCTANGGCANNTTTGANAACCTCCGGCCTGTTTAAAAGCCNATTTT

SEQ ID NO: 463 ACTTTTTTTTTTTTTTTTTTTTTTTTGNACACATTTNAAGGGTTNATTTANANA  
AATTNTATGTTAANCCNTTGAAAATGAGGAAAAGATGGTTNACAAAACCCAAGATTNAATGANGN  
TTGAANCNGNTAAGGCAAAGTNTTTCTTTGGTTCCGGAACACCCGGAAGAGCCATATGCATCCTC  
TGGAACCANCTTTNTTACTAATGGCATCTTNCACANACTTNGGNAANAAGCACACGGCACTTN  
GAAANTNAANTTGAAAAANAANNNTTNAANAGGCNCATTNCTTGGGGNTGATNCCCGGGTNN  
CCACANNTTCTTAGGGNNGANACCTNCATGANCNCTGATNCGACTTTTNTCTTNGGAGGGGGTG  
GAAACACCAGGGGACCCCTATGACACTGGAGNCGAAACATGCATCCATTGGCNAACANNGGAT  
NCTGGTGATNNNAGGGAACCATAGTTCGGCCTGTTATGGGGGCTCANCCTGGAGCACTGGAANN  
GGGNTACAGGCTTTTTGNGTGACTGNTCACCANTGGGNAACCACAGGGACTNGGNTCCATGAAAA  
CTTATGTCAANGTCANTTTTCCCCCGNTTANCTTGGCCNGGGACACGCTTANGGGCNAATTCAG  
GCNTTTGGGGGGCCGTTCTATTTTGATCCCCGANCCTCNGGTACCAAGCCTTGGGCCNAATTCCTG  
GGCATTAGNAGGACCCCTGNGTTNAAANTTGAATTTTCCCCCTTNCATTTNCCCCCACAAANATT  
ANAAANCCCGGTNAACATTAANNTTGGTAAAGCCTTGGGGGNGCCCTTAAANGGAGGGGGGCC  
TTTACNNNNCTTTTTTATTGGGGTNGGNCCTNACTNNGCCGGGTTTTTCCAANAACNGGGGAAACT  
TTGTTNTNGCCAACCTTGGNTTTTTTGGAAATTNGGNCAACCNTNCGGGGGGAAAAGGCCCNTT  
TNCCTTTTTTGGGGCGCCTCTTTCCCCCTTTGGNGNGAANTAANTGAAAACCCNTTGGGGGGNGG  
GGGGCG

SEQ ID NO: 464 GGACCTCCCGNTCANNGCTNNTCATNNCCNCTCNCNCNGCCTGGTGGANCA  
NGCTGACCNNATCCTCTTNTGNATGGAGGCNCTNTCNGGGANGGGGAACNNACCATCAANTNA  
NGGNNAANANGGGGTGCTAANGNGCCNCGGNGCACGCTCNTGNANAAGCNCCNTAATGAAAACC  
TTGTNAGACCTGCGCACTNACATCTCCCTCCCTTNTCTTCTCTNANGTGGNGGAGAACCACAGNTT  
NCNGAGACANGCAGGCTGCCNCCAGNNTGAGTTNNNTNAAANTTGCCTTGANNNTGAAACNTCCT  
TTCAATCTCCTCATTGATGATGCAGTACTGNCCTANGNACTGTGACCCATNCACNAGTGGNCTTT  
NGAAGNNNCCTGTTTCTTNTCTNCTNTNTTNCGANGANCCTNTNCCCTTCATTTCCGCCGNCCCTN  
ATAACACGGAGAAAAAACNAAGCNNGGTCCCCAGTGTGATGNGCCTTNTCAGNCACTAACTGTTT  
NCNTGTACCNGACCCGCTCNATNGAAAGGTTCAATNCAACACACTGGNGGCNATTCTAATGGNC  
CCN

SEQ ID NO: 465 GNGGNGGCCNGGCCGANGNGCACACTTTGGATACACTGGATGCTCATGTCAA  
ANGGGGTCAACTCATCTTCACTCTGAGATNCAAACNTAACNCTTGGCGGCATCAACCAAAAAAAT  
CAAACATCTNTTCCNGAATATTATAGNCTCCACTNGCTTNNAGGGTTTGTNTGGTTNTTNCNN  
NGCTTCTTNATCNCCTCNGTTTTGTGNTNTGACTCCCACTGACATGNAGGCTACTGCTCACTTT  
GGTGATGNAAACGACCAACTTGGACANAAAAACCCCGCGNTCTNCCNGGGCNGNCGNTCGAAA  
GGGCTAATTNCAACTNACTGAAGGNCGTTNCTACCGGATNNGANCTCGGTGCCCCCATCGGCATT  
NATCATGGGCATGGCTGNTNACCTGTGTGAAATTGNTTACGGTCATNATGCCNCAACATACTAT  
GCGGGNAGCATAAAATGTCTAAGCCTTGGGNGCCTAATGANTNAAGCTTACTTAGATTGAATNGC  
GTAGNNCTCACTGTCCGTTNCCANAGGGGAAACCCNTTGTGCNNGCTGTATTTNCAATCCGGTCA  
NCCCCGGGG

SEQ ID NO: 466 GGTACCAAANCNANTNACAGGANGGGCGGGACTACCGGAACCTACAGGCTGT  
TATCCCTTCCCGAATTTGGAATTTTGTANGAAATTANTGTTTCTATGGAAGAACTGCATAAATTA  
CAAGAACNTNAAATATNNCACTTAATTCGGNCATTGGATNCNNGNTGGATGATAGCATNTGGCTG  
AGCCNCTGGTTTGTNNTANAGGAANTGGATNNCTANNGGCCNTGAANGANNGANNTCCTNTC

ATAACNCATGCCACNGTAGTCCTCACTGGAATTGGAGCANATGAANANCTTGCACGNTATTTTCG  
 NNANCCCNNTTCNGTTTCNCNCANNTNGNCTGTGACAGAATGNNTATGGAAANAATGATGGAANTG  
 GGTCCAATTANTTNTCNAAATNTNGGTNGNGTNGACACAGATATTNTNGATCATGGATAAGACNC  
 TAATGTTCCNTNCTGGGTGGACNTGTGCTGCCTTATTAATCTCTNGNAGTNGGANNAAAGGGACTGA  
 CTTACNCCCAGGATTGGGGAAAAAGGCCTNTCGCTTCCGGGGGGACT

SEQ ID NO: 467 CNGCGGCCNNNCCGNCGGGCTCNCCCNTTNCCCAGAAAAGCGGNACTTGCTG  
 CNAGGGNNNAANGNCCAAAGNNNTNGTGCCTGCGTGNACNGACACCCTTGAAACGATNTATGTA  
 TNATNGCTACAGGCATGNCNGCCANNTTANACNCCANAGGAANGCTGNNTNCNCTNTNTNNGAC  
 CAGCGCCNNANTATCCTTCCACNGATAAGGCGTGCANCTTTTGTCTCTACCAGATNATGANCTGNT  
 TATNNNTTGATAAAATGNNCNTNTANAAGTGGATGANCANCATTGAANGGANNNNTGNTTNG  
 GGATATNGANGGACNTNNNTTAAAGANTNNNAANACGNTGNNTACCACANNNNAAATNGGCGA  
 GATNTGCTNGNGCTCGTCTACTGATGACNNNTGTTACCAGTGGATNTTTGTCAANGATCAAAG  
 ANTGGCGGGACGGGACGAANGATCTACCAACGTAGGCNCCCCACCCTATNGGNGAACTCNAATTC  
 ANTGGCGGNCNTTCAAGAGGACNCCAAANCNTTCCAGCTAGGCACAAGCTGGTATNTCCCTGNT  
 GCGGGANCTTTATCCCANAAAATTCACCCCNNTNTACTCCAGGCTNANGGTAAAGGGGGGGGTC  
 CCNNNA

SEQ ID NO: 468 CGCNNNNNCCNNNANGACCTATNCATGCNTATGNNGNATGTNGTTGTGATNN  
 TANTACNNNTCCNANCATGNGNNGCAGCNGTNGNGANCATACTNGGCTTTTTCNNNNCTATG  
 GANCCCTGGANGGAAGGATCCACACCCTAAGCANGGAGGNGCTNNGCGANTTNTNNTANAGAC  
 CTGNNNAGNTTGCNNCTNAANAGNNAANNNAATGNTGCCAAGCATCAANTGTNCATGGACATGN  
 TCGCGANNATANATAACCACTGGNNTCCNTTNTAATATNGCNTCTTCGGTGCAANNATGGGTG  
 NCANNACATAATNCTNNGGACGNTGTGGCAANTAGGANTACTGGTGCCCNNTAGGCGNATGNTGC  
 ACCCGAGTTCGANGAAGACCCACTGCTCTGGCAACTTNTGNGATTGATTTATACTCCATTACACTCT  
 TTTAAAACTGGCTGTTTCCCACTGGGTAACTCTGAAGAACAANGCTCNTGGTNGAATAGCTNGGCC  
 NAAGTTTAAANGCACTTCTTTATTTGGATGTNTTCNGCTTACTATGGTAATACCAACCTTCCATGG  
 GGNTTTGGGCTANTTGGAGCGAAATTCCTTTGTCAANATNTNGGTAANGGGAATACAGN

SEQ ID NO: 469 TTTTTTTTNTNGNTNCTATNTTTTCTNNGGCGNACCGTGCACCTTTATTGAA  
 TGACACTGANGACAGGNGNNNGGNTAAACNGNTGCTTCTAGGGGACGACCAAGGGGGCAGG  
 GGCTNCAGCCCNAAACGTNCAGGGCCTGCATTGCACAANGNNGATGCANANGTTGCNAGNTNTGG  
 GCAGGAACTAAGGAACCCNCTTGACCCGTATTATCTNAAACATAGANTTGGTAGGACTGNTNCA  
 TAACCNANNCCATCANCCCGANNANTCCCATGGTNATGANGNTGCCCAAAATCNGGGCCCATNA  
 GTTCNCGCACTTTGCGGTGGNGGCATTAANCCCNNGNCCCCANTAAACNTTCCCACCATNTTNTGT  
 NCCCTGANTTNNNTGGAGAAAGNAAATGNTTNAANATCCTAACANAAGNNTANCAAAAATGAAT  
 GGGTGNTAACANGGACCAAAATNATGNTNGGNTGACTTCAAGGTCTNATTTNGGANNNTTNTCT  
 ATNGNANNNNCTTTNAAAGGNTGNGCTTNGGNGGGACCCTCATANNNAACCTTAANTTTTGCCT  
 AAGACTTTTTCTNNGGCTTGGGGAGGNNANCNTNGGATAANNNTAANAAGGNGATTTTGGANAT  
 ANAAGGCNTCATTGTTN

SEQ ID NO: 470 GGTACNNTNTCANNNNATAAGNGCTGNNGNNNNCNANAATGANGGNGTCA  
 TNGNNTGNGGTGATNNCCATCNCGGNAGCTCAGTCNNAANANATTATGGANGANAGGCTTTGN  
 NNGGAGTGATNANCATNANCTGGGACNAACTGAGCATCTCTACTCANANNTAATACCGTGAGANA  
 GGCACNTANGCACNAGCTTGTNTNACATATGTCCAAGGCTGTAGGANNTTGAATCTGTGTAATC  
 AACGTGAATNTGCTNAATAGTAATGANAGCTTGTGAATATTAGGCNTAAAGGTGNTGGTNAAATNT  
 NNANNCNACTGCTCCATNTGAATTATCTACTGNGTGTTACTTACTGAAGNTGAAATGTNNCNNTA  
 TGANATATNGTTNNANTGNGCGTTCGANGANGATNNGCTTTTGANNNNNATGTACTNNGTGCNAC  
 TGAANACNNNNNTCTACNCTGATNTTNCAAAANNGGACCTGCGGAGGATCACNCTTTANACTTT  
 ATCTTTATGGGGATCCTGNGACCTNGANNTTGGAAAAANATAATTACAGTCCCTGGCCTAGGAC  
 CCATGATANCTATAATGGGTTCCNCTTCTTNNCCGTNNNCNCTCNAAGGCNAATCCATNCCTGGC  
 GGGCGTTCT

SEQ ID NO: 471 GCGAGGTACCACNATCACCAACCAATACAAGTTTGAAGTGGACCTGGGGCTC  
 TCCCTGCGCAGCCATCGCGGCGTTCCTACCAGAAGCGTTTCCGCGCCCGNCCGACTGAGGCTGGA  
 AAGATTNNGGAAGCGCCCTNTACNTGCCCCGGGCGGCGNGTTCGAATGGTCTNATNTNATTACACTT  
 NGCATNCTGTINNNTNGATNNNANCTTTGNACTNAGTTGTTNTGTNATTGANGGTACATNNCTT  
 NCNTCTTTGNGNTNTNTTGGTNATTCATCANNANTNNNTCTNNNNNTTATTNNNTCTTTATNATNTT  
 ANTTAATTCNNTTTTTTTCNANATATTTGAGTTGGTATTAATNGTNATNTNTNCTTTNTTNAATTT  
 NATATTTAANTTCTTTATTTGTTTTCTATTTTNTTTTNTTTATCNTTTNTNTNTNCCTTTTNA  
 AGNTTTTNTTATANGATNTNNNGINTTACTGTTTNTTNTTNTTCCCACTGNTNGTGCCTNTNTNTN  
 TATGCTGTATTTTNAAGTTNTATNTTGTCTCTATTTTATNAATNTTGTCTAATCTTGTCTGTA

TNTTTTNTTAATTNNTNTTCGTTNTAATTTATTTANTTNTATTTTCTACACTTTNTNGTNTNNTATGT  
ATCTTANTTTTTCTNTTGTNTATTATATNAAATATNATCTATATTGATTGTTCTATCCNCTTTTTTTNG

SEQ ID NO: 472 GTACGCGGGGGTGGGTGGGATTGAGGTGTGCCCTNCNTCATAAATACAGACT  
CANTTTTGGCTTGCACACTCTGAATCTTNNACCGCATNCTANCCNACGACTNACACAAAGCACGTGN  
GTGAGGAANTCCAGANTTGGCATGGAAAAAATANNATTGTTATCATNATTNCNCCTTGTGAGCC  
CTTNCCTACACTCTGGCCAAAATTTCCATATTANAAACCTTANCCANAAAGGACAAATATGNCTCT  
CNGACNCATTCTGNCCNANACCCNCTNCAGANGTTGNNGTGACCAANTAATCTNGACTCANNACA  
TTTNAAAAAACCTCTATATAAANCCAATANNANCANNAANNCCCTGGANNATCANTTATTGA

SEQ ID NO: 473 TAGGTCCNGCCGACGTACACTCGAAACCAAATNNCTAAAACTTGGTTNGCNT  
AAAAAATNGTNGTTGTAACATTAAACCATAACCTAATCAGTGTGTNCACTATGCTTNCACACTANC  
CAGTNTNCTCACACTNNTTCTGGTTTCAAGNCTCAANGCCNGACANNCANAAAGGGCTTGGAGATT  
TTTNCCTTACAGTACAANTCTATCAGCAACTNTGAGAGCNTNNTTNTATGTNGGTCANGCAACAG  
ATCTGTNTCTGCAAGG

SEQ ID NO: 474 GCGGNGACGACNCGNNCCAACGTGTGCCCTATNAACTCTCCATGGNAATCC  
CCNCGCCTACCATGGTGNCCACCNGGTGACGGGGANTAAAGGNTCATTCCCGATANGGAGCCNTA  
NAAACCGCTACCNGATCACAGNNGAANGNAGTNTNGCGCTCATATNTTCCACNCNNGACCNNN  
NGAGGTAGNTNACNACTGAGTACNTGNNNACNTAANNTNTGGCTCTTTTGNCTTACNCNCAAATG  
GCGAAGTGNCACNNACTGGNGGNCGTNTNTNGATNCANGCTNGGATCNANTCTTGCNTATATN  
GANNTNATTGCCTTTTTCTTTGNAATAT

SEQ ID NO: 475 CGTGCANTTNANTGCATAAAAAGGCCTCTCTCCATNANACTCANCACTTTAC  
AGATGTANAATNTATAAGCATGCCAAATNGTACTTATCTGCCACATACAAAGCNTCATNCCANGT  
GCTAGNGAGGGGAAAAAANGTANGAGATNTGGCCCTCAAGGANCAACAGATNTTAATCTACCT  
AACNAAGTCCNNAGNGTNNNCCAGGCATGNAAAAATTAGTGNTGCTACATGGAT

SEQ ID NO: 476 ACNTCTTTGACATTTTCAAANTGAANAACCTGNCNNNTTTCATTTGNAANGG  
NNCCTTNTGCACCNACTNGNNCAAGTGATAGAAAGGTGGNGAGACAGGATTGGATCACAATANG  
ACATAACCCAGAAGGNTNCNTTNTNATATTCTTAAACAAATANCNGNTAACGCCGGAATTTG  
TTGNTTCAANGGCTTCTTNCNCGGCCTGGAAGTTANAGGAATTCTNGTAAAANGCGGNANANAG  
AAACANCCCCCAAGGAACCTCAAAGGCCTAANTNTGGAAGGAACCCGGNCNATGGCNGGATCAA  
GGCNNTTNTNCTCCGCCGACCAACCCCTNTGGGGGATNATTNCACCTGGNGGCCGNCCTACTGGT  
GNTACCTCNTTNCACCTNGNGGAATATGGTNGTAAACNGTTGCTTGCNNGAAAAANTATCCNCNCC  
AANANCCNCAACATCTANCNGNAGGATCATTGTANNCTGCGNGCTNCTAATGAGCTACTACATT  
ATTTGTNGG

SEQ ID NO: 477 CCGNCGGGCNCCTGCCAGCANCGGACCCTCANAAGAAANCNCATNACNTCAG  
ACTGGTTCTTTCTCCATAGCTCCTGGATGTNCAAGTATGTATCATGATTNTTNACACNNAATNAATN  
TTNACCAATTCAAATNCCGNAGCTATATGAGTATTCTNATAACCAAGAATACACTACACCANCT  
NATGACTGNTAAAAAAGATGNANTCCNGCCTTNTTGGGACNGAAATTCATATNNTACA  
AAGNNGGCNNCACCCCTCTANATAAAGCGTTGANGNAGGTCTGATTTAAATGCTGCATTCCCTT  
AATGCTCNTTTAAATGAAACACACANCCCCNGAATACANTATGTTTTGAAAAGAATTCNTTAGC  
ACATCTATACCATTAAATATTANCAAAGNGTATTGNCAATGAATNAATGNGTCCTTATAGGGCCT  
NNGNCGNA

SEQ ID NO: 478 ACTTTTTTTTTTTTTTTTTTTTTTTTTTCCGNTTTTTTTTTTTTTTTTTTTT  
TTTTTTTTTTTAAAGGNTTGTGCTTTANACNCAGNAATNATTTNTTTANATACTAANCANTTA  
CAANATTACCCNTNTAAANNAATCCATTTTTTAAAGNCTCATAAGGNAAAAANAACCTTCATAAT  
NTTNTNCTTTTANATTACCCAGTCTGATAAACNTAAACANCTTAAACTNTGGTAANAGTTTAA  
ANAAAATTTANAAAANAGGCCCTTTNAANTCACCTNCTGNNTNTAACTTNANANTTAANGTTACT  
TGNANATNTTTCATTTNATNAAAGGTTNTCCCCANNNAAACCAAATTTTNTTNTATGCAAAACCN  
CAANAAATTTNNGTTCCAAACACTNTTTTNTGGTTGAAANANNTGCAAAANAAACCCCTAANTTA  
AAAAAATNTTTGAAAAAATNAAAAACAAAAACCCCNAAATTGAAANCATTTACAANCCCTT  
NAANAAGGCCTTCAATGTAAAAANAAGGTTTAAACCCCTNNTTTTTTTCAGGTAANGGANTNA  
AAAGNCCCAAAAAAANTCNTCTNCCNANGGGGAAAGANACTGGAAGGTTGGCATTACANAT  
T

SEQ ID NO: 479 CGCGGCGAGGTACNCGGGGGAGTGAAGGGTCTGCTGCTGAAATTTGGGGGC  
AAATAACCGNAGTANGTTTGTCTGTGCCTTGGNGAGTCNCCGGCCTACTGGGAACGGGACTTNT  
AAAANGAATATGTCTNGAANGCTGTGGTCTAGGCCATTTTGTGCTGCTNTAAGCGNANGTCTCC

NNANCNATTTGGAGCACACATCNCNTNTTTAANTGGAAGGNAGTTTTCTCCCGAAAANAACTTA  
ATTNTNTTTNNGCTNGANANTNCNCTTATTANATTTAATAANATTAGAAC

SEQ ID NO: 480 ACTCGGGACTGTGTCANGAATGATGGAAACCAAATCATGGAGAGTTTNNTTG  
NCAATATCNTTCNGCAANANAANAATTCGGGTACAGATTCTTATNAGNTGANATTGGCATGTGT  
TTNGTTACACT

SEQ ID NO: 481 ACGNGGTCGAGGTACACANTGGGGGCTCCTCATACATGGCCTCACATTGAGG  
CAACCTTAACTTTGNCTAANANACTTCATCTTNATACNTNCTCGCCNTGAGAAGANNNTGACCTT  
ANATTNGATGATGGGGCAGCCNGTGGGGTANTAAATCGTGAA

SEQ ID NO: 482 GTGTGTGTGTATGTGTGTGGTGTGTGTGTTTTAAGTTTANCCTTTTGTTTTT  
GNTTTTTGGTTGGCANNAACCGATTTTAATGACTAGCTTTTAAAAATACAGTNCAGCAANTAGCTG  
CTGAATATTGTNTAGCATGTCNCTTGAANCTACTTGNATCTACGGNGGNGCTCCTAANGGACCACA  
ANTNCNINATTTNGGANAGAGGTGTGGAAATCTTGATTGTCNACNCCTGGAANACTGNTTGACTTT  
TTAGAGTGGNGGAAAAACCAATCTGNGANGNACAANCTTCNTTNGNCCNTTNATNAACCATATAA  
ATGCTTNANNTGGNCTTTNNANCCCAATTNNGNAAAAATGTNGGATCANATANAACCTTATTGGNN  
ANCTTCAATGNNGGAGNNNNNTGGCCATAATAAATTANANNGTTNNNTAAATTTNCTNAAAANTT  
NTTNNTTAACTTTGGNCNCTTGAAGTGNNTCCATNCCAAACATTTTCAAATGGGGTTGGGGCNTTA  
NCTNAAAANGTNGAAAACNTTTCCTCTATGGAAAA

SEQ ID NO: 483 ACNTCCNTCTTCCAACTGCTTGCCAGCAAAGATCATNCTCTGCTGATCAGGGA  
GGAATGGCTTCCTTATCCTGAATCTAGGCCCTTCATTTTCTATCGANTCCGANGGTNNAACNCNG  
GGGTGATGGTCTTCCNTANGGTTTTACCANAAATCTGCATTNGGGGGGGCTCCACCTCAANGG  
CCGTATCGAAAGGGCCNNNNANNNAAACANTTGATACCCCTNCCTATCCAAAAATAAGTTTAACT  
GNNGCNTTGAAAAATTNCCAGANNTGGNGGGGGAACCTTAAANATTCCTCCAAAAAGCTNGAAAG  
AACCAANTAGATGTNANCTGGGGAAANTAACNAANCCCCCAATGCCANNGAAAACCTAAACNCC  
TTGGCNTTAATTTAACCCNGTGGGGCAACC

SEQ ID NO: 484 ACTTNNTTTTTTTTTTTTTTTTTTTTNTNAAGGGTGATAACGCTCTTTTCANAN  
ATCATAGCACATGAAGAACCCATGGACACTACACAGACTATGAACAGTTACCAAAAAAAACTCG  
TGACTAAAGNGGGGATTANCNNCAAAAAAAATTTCCAAAGGNGAAAAATTGGNAAATCNGGA  
ATTTTTACNAAAANGTTAAAAAATTATTTATCTTTCANCCAAATGAGGCCCTTCNCATTCCCAAAA  
NCNTTGCTTTCTCTCCTTTTCGTTCTTGCNCGCGTTTCTTCCGNTGCTNTTCCTTTGTCTGGAAANGG  
GGTTNTTTGTAAATTTTNTTGAAAANTTAATTCTGNTGACATTNTGANTTAAAAAANTATTTCTTT  
AAANTGCCCCCAAGGNTTTAACTTCCCTTNGGNGGGTCATANNTGGNNGGGACAAAAAACTTTN  
TTTAAAAAANNAACCNNTNGTTTTTTTTNTTNNAAACCGGGATNNNCCTTCCCCCAATTTGGCCNC  
CTTNGAGNGGGGAAANAANNCCNCAAAAAATCCCCCGNGGAGGGGGTTTNAATTNGGGGGG  
CCCCAAACNCCANCTCCNCTTTTGGGGGGGCCCCCNAAANNCGGNNAATTTTCNAAAATT  
TTTG

SEQ ID NO: 485 TGCCTAATNACAACATGGATGACTNGCAAAGGANGGGCTCTTTACTTTAAGC  
ACCATANAAAAAATCNAACGCNCANATGGNTNATTGNNNTCAGNTATGNCNCTGAATNGAACNTT  
TGGCANTAGGANTGAAANCATTTGGNAATNTGNCATTAAACATGTNTNATAAAAGGCNTNGNGTCN  
AAGGAATNNACCCNCCATNNTTGAAANGCNGTTTTNCAACTAGGCTTAAANCACTCTACNCTTTT  
CGTCTGTGTCCATCNTTGNAAANACTGGCAGCNGTTTGNANNTGAANNAGNATGAAGAAAAAGAT  
ATGGGAGCTTTTNCCTTTTNTTCTGTTTAAATCAAACAAAAACAAANGTCAATTGGTN  
TCCACNACATNTTAAAGCCNATTGGTTTGGACCTTGGGCCGGGACCACNNCTTANGGGNGAAT

SEQ ID NO: 486 ACTTTTAAGAAAAAAGCAGGGCCTTGGAAGTTTGGGTCTTNTTCTCTCC  
CTGTTGCAAAATTTCTATGGNTTGGGTTGGGTGGCGGNAACCGCCGNTTCNTNCTNNANGGTGGC  
ACTGCCACGGTGGGCNGGCGGTCCCTCTCTANTCTAANGGGGACCCACCGTNTANATTCTGNAA  
CTGGAAGTGTGNAAGGTGAATANGNTCANGNNGGNCNTTNTNTTTTANNTCAANCTNTNNTCTT  
NTCGTGACAAAGCATANCT

SEQ ID NO: 487 ACTNATNNTGNTNNCCCAATNGATACTGNTTGNACTAACATCCACTCTNGTAT  
GTNGCTGAGATAACTTACTTTGACTGNCATTTGGATATCTCTCANACANACCTTGAACNANATAA  
AAAGGACANGTTTGTGGACTGTGT

SEQ ID NO: 488 ACCCATGCTCACACNCACACACTTCCAGTTTTATACAGAATTTTTTTAANGGA  
AAGAAACCAACCCAAAGTATTGCATTTGAGGNGACACTCCCTGAAGANTTNTATACAGAGGTANT

NTACNNGTTTACCNCAGCTGAAGGTTCTTTTTATTCAATTCTNTTTTTAAAGTGANCCCATGATTTGG  
TGNCNTTCCNAGATTATCCCCTTNGCNACNACNCAAGCAATGATATTTAACNNAGGTNANGCATA  
NTCNAACTNCATATAGAGAAAAAATCGATNACGAAAAATNCNTGNAATAAGACTNAATNACTTGGC  
CNTAACAAAGGNCTTAGTAGATATNTAA

SEQ ID NO: 489 CTCNGNGAGGCCCAACGCTGCTNNTGCTACACTATNGNCAGGATNAATCCG  
GTCCNGTGCGGCANATGGNTANTCTCCTCTGGCCTTGGCNGCCCTCCTTGTTCTGTTGGACANGGGAN  
ATNTCNNGTGGCANNAGGAAAGCTCCNTTTTTTCANGANCTGCNCCA

SEQ ID NO: 490 ACACNCGCNNNCTGGNNTNANNACGCAAGCTNANNATAGGTCNGGNNAGCG  
GCNCTCTGNNANTGGGTNNNTGGTGCAACNNTNTCNCAGGCANGNGATGGTGCAATTCCTCGCACT  
TNGCTGGATAATGAAATCNAGGATGACGATGCCTTNTTCTGCTGTTCTAGGGCATCCANGACAAT  
GCCAATCGNTGTCTTTTCAAACCTATACCTTCCATTTTATNACTCCAAGNTAGTACTTGAGGTNCC  
NAAGTCTCANCGATGCNTGGATAAACATCCACNGCNAANGAAAAAGTAATGCCAAANANAAANG  
ACTATGANTCNTGANTACANGNGAAATAAANANTTTTNTCTCNCCAACNTGACCACCATAGCGANC  
TTTAAANGAGNAACTNCNTTATNCCNNCNAAGGACGTCNTNATNGTANAGNTGCNCTCCCNNA  
ACNACTGTGAACACNGANCNGGANCAAGAATCCTCCCAAGGATCATN

SEQ ID NO: 491 ACCACAAGGATGTGAAGCATATGAACTCTGCANGANTCCTGNCCACACTGAG  
GAATTATGACTACTATGCATGCCNAGNTCCANAATGTATTAATAAATGCACTTGCTTCCTTANAGGA  
NAGTCTCATTTGGNCGNANGCGNGTGGCTCATNCNTGTANTCCNNGCACTNTGGGAGGCCGTAA  
GCNNGTGGTNNATCTTTTGCTCAGGAGTNTNAGACTNTCCCGNGCANTATCNNGNANATCNCA  
CTTTACANANTANNANTANAATTGGGCATCGNGTATGCCCCAG

SEQ ID NO: 492 GTACANAGTAGCCGTGATGTGGTCATTNGTCCTGATGCCAGCCTTGAAGATG  
CAAAAAAAGAGGGACCATATGATGTGGTGGGACNACCANGAGGCGATCTGGNCGCNCCNGANTC  
TATCTGANTCTGCTGNTGTNNANGANATACTGNNGCAACNNGAATACCCGAAGNNGCCTNGATAT  
CCGCCATCTINTGCTGGTTCTACTGCTNTGCTNGGCTCNTNAAANTANNTTTTGGANNCAANGTT  
TCAACACACCCCTTTTGCTANAGACAAAATGATNAATGGAGGTNATTNCACCTACTNTGANAATCN  
TGTGGAAAAAGACNNTCTTATTCTTACACTGCCNGGGGCTGGGACCNTCTNTAGTTTNTCTCTT  
NCNATTTNTGAANCCTCTGTANTGGCANTTGAATGTNGGCGGTCAAAGANAANGCTCCACTTN  
NTCCTTAAANAACNNCACNCAGNGAAATTTGACTCATTNCCTTANTAATAAACNGGGCNNCTA  
ATGAANCC

SEQ ID NO: 493 CNGTGGNCGCGGCCGAGGNACAGCATCANTGAAAAACACANTGTCATGAAA  
CACANATGCTGTNGCATGATGACAGTCACAGANCCATGCTNAATTGCTCAAAGAAGCTCCCANCA  
TNCTACCCAACGGCNGGAGTANAATTGNTNNTNNTNANAAATTTACAAACTTCAATCTTTTTTAC  
ATGGTATTTTACCTTTTNTGAANNTCATTACANGGAGTTNGTATAANTATTAGGCGGNNTNTNCC  
TGNCAGCTCATNGGAACCCAGNGGATGNGATNCCACCCCANCGAGGNGGTTTTNTGTTGTCCGG  
CNGGATCATATTCTTGGAAATAGGAGCATCTCCANTGGCCCCAATTCTGGNGAAAGGAGAACCCC  
TTCAACCCCCCATATTTTACAGGTTGATTTGGAGACNNAANGCGGGGCACCTTTTCNCTCTNAAGG  
AGNNANGCATCCCCNGNGGGTGTCTCATGGGATGGNTANAGGGCNTNGNGTGAAAACCTCNATN  
CCANAATTCCTATTGGGGGNNTGGCCATNTTTCNCATGCNCCNAAANAA

SEQ ID NO: 494 GGNTCNCGGGGAANAGGAGTTGGANTATGGGGGACGCGGNAGGCGGCNTAN  
ACAATNAGTTTTCTAGGNTGCTTTTTNGNTCCANTTTGGGAGATCGATATTTGTTCTTAACGATGG  
GGAAACCANTATAANGGCANAANTGAATACNTGANGACGGCTTAGGGGAANTGNACTTNTNNT  
TTTGACGGANANNCCGCNTCAAGGAAAGGTAAACCTCGCNTTTAAGNCACTTNAAGAGGCTAN  
GAACACCATNGGAATTAGNATCTGAANTNNGCANGGNTGGATCGTNACNTTNNAGTGNCCTCGAG  
TCNTNCCATTGNTAATAAAACCNATTNATAAGAACTNGCCCTACCATGNTGANCTNTCTCCNT  
ACNNTANTTATNNTTTTAAATTTATGGCTNTCANNTANGGCCNTATTGAANCTTANCGTNCGNTGT  
CTATNTCGCANTNGATGGCATATTCGTTTAAANANATCAATTNTTAAANATNNNNNANTNNATTN  
ACNENNTCGANATNANCTNCANTGGNTCNTCANNTCNNCNTNTTCTTGNGTNAACAACCTTT  
ACTTAATATGAGATTGGGCNTCNAACANNNTTCCCTATGAAAATTNNAACNTACTACACNCACA  
NNNANATTNTANGNTCNANTGTCGCGTATAANTATCTCTACNTNTGCTNAATTAAACCCNCNNT  
NNTGANNCCGCGNCCG

SEQ ID NO: 495 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGGCTCAGATAAATTATTTATTATATTC  
TGGATTATCCNTGGAATTTTTCTGGATATGAATAAATAACAATGTNTCTCANAAAAACACATTTTG  
AAAAGCCTTATAAAGTTAAACCTTTAAATGCTTTTCATCAAAATGTTATGACCAATAAAAAATGCCA  
TTCATGTATCTCATCTCACTTTCTATACATTTTGGAGNGATATGTATTTATACATACTGCAGTTGGA  
AATAGGGAATGCTTTTTTGGTTTCCACGGTAGGCTTGACAGGGTTTAAAAGCANAAGTCCTGAANC



TNTNAAGGCAACATAATGAAAAGTATTNNTTTTGAAACCCTGGAGTTTTNANAATATATANNNAT  
 TTTNTGGGNGAGGGCCANTNTNTGTGGCTCANNCCTGTATCCCNACNCTTTAAAAGNTTNAAGCNA  
 AAGANTTTTTNGANAANNATAAGTTCGANNANAANNNTGGGCATCNAAAGNNGNCCNCATCTTTT  
 CNTAAAAAANAANTNCNTGANCCCATGGGCNNCCCCGGANNCCCCACTCNTANAGNNGNTNG  
 AGGGGAANATTTGTTGNNCCNCNAANNGANGCGGCNAAANCCTNAAGNANNCTNNTNTCCACC  
 CTGGAANAAGNGGCCTTGNTTCNAAAAAAAG

SEQ ID NO: 496 GCGTGGNCGCGGCCGAGGTNCAAGATTTACCAGAAAGAGAGTGGTGTGTAN  
 ACATGCCTGGAGCAGACACCTTGGAGCCGCTGACAGAAGGTGAAGCATTCAAGAAAAATGTGGA  
 AACTTTTCCGNTGCTCTACACNGTCNANAAACCTGTCCATTAAATTTTCGNTGAAGCTTTGTCTGATA  
 GATAANCAAAATAGACAGTNNAAAGTAAGTTATCTCAGCNACATTTGGGGAGTGGATGCTGNTGAAT  
 TGCGATTTATTTGGGGGAGCCNTATATGTANCTGAGAGGTGGGAATTTTATAGATNTAATNTAAGCC  
 ATGNNGGCCAAGTINCATGANTGCATCAATGNTGTTATNNTGGGAACANGTGAATTTCTATNNNT  
 GTGGGCTNGTTNGNNGCTTTNTTGATAAAAANTNAAGGCCCCNGTTTACTNTATNNGTTCAATTTG  
 CAATCNTTGGTTTNCCTTCNTTTNNAATAANNATCNTCACNNNATTCGGNGCTCGTGGGATTTNTC  
 AACCTTAATAACCTGNAANAAATAATNTGCTGTGCTAATTAATNATCNTAGCNGTGNTTANTNGTN  
 NGCAGCCTAATGGACTATNCNTANCCATTTATCTNATTAAGTGTTATTAGTGATCTCTTTTGATNN  
 TGGNATGGCCCTTCGTGCTATATTNNTATNTTAAATNCTNGGTTATNANTGGNTT

SEQ ID NO: 497 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTGGGGTTTTTTTTTTTTTTTTTTTT  
 TTTTTTTTTTTTTTTTTTGNNTCCCNCTACTGGTGGGGNTNGAACTAAAANTCANGCCTNTGCCT  
 TTTAAGCTNAAGGNCGNTGNCCATNTGCTTTTACGNTGCCTNCATNANCTTNCANGGGNTTAA  
 AAACCCCCNTNCCAAAAAANCCAGGGNNGGANCNACNCTNGNGANGAANNANGGNNCCCCNC  
 AACTTTTNNNTGGGTTTGANGGGGCNNTTTGGCTTNNCTNTTTAACACCCTTTTTNGAAATNGCCT  
 GTTNACTTTTTTAAAGGTCAANNGNCAATTCCTGNGGGGCNTTTTTTTTAAACCCGAATTTGGNCT  
 TTNNCNTGNGGGNATAAGGGTTATCNAAGGNNTTTAAANAAACCCCCCGCAAGGCCCTTNNGGG  
 AAAAAATNTTTTTCCCCCTTTCCAAAAGGGAAANAGGTAAANGNNGGNCTTNTCCGGGCGGCCTT  
 TTNAAGGGCAANTTCAACNANCTGNNGGCGTTTTTAANGGGATCCCACNNGGNCCAANCTTGG  
 NGNAATAANGGNNTACNTGGTCNCGNGNAAAAATNTTCCGCTCNAATTCCNNAATAACATNCGG  
 AAATATAAAGGTAACCNGGGGCTNAG

SEQ ID NO: 498 ACGCGGGNNGTNGNAGCCTGNGGGNCCTANTGNNNNATNGATNGNATNAT  
 CAANTATACAATAACCTTTGNGNNTGTGNCNAANTGGAGNGAATTTGNNTTNCNNGGTATCATA  
 TGCNCTNCAGCATANGGNAANCTGNAAAGAGCTATGCCANANGACTATTAGANTGGCAAGCCTC  
 TGATCANAGCGTTANAGATNACGAGANANANGGGTGTNATNNGTGCCTNNGCTGTGCCNTNCATG  
 GANNGANNNCCANAGNGGANNTGTGTTTATGACTTGGCAACTCAATTGCATTGGAAGCAATAAAA  
 CCGACTTATCAAGTCCTAAATGTACC

SEQ ID NO: 499 CATTGANCTCCATAGAGACAGNGCCGGGGCAAGTGAGAGCCGGACGGGCAC  
 TGGGCTGACNCTGTGCCTCGCTGANGAAGAATANTTTAATCGTGGGCAAAGGANATCCTAANAAG  
 CCAANAGGCNTAATGTCNTCATCTGCTTTTATTNGTNCAATCTGTGTNTGAGAAAACCTTNNGAAN  
 ATNCGCCCNNTAANNCAACGTGANGAATTTCTGAA

SEQ ID NO: 500 ACCNNGGNCNNNGCCGACGTGCTAACATGCTTNACNNATCANTATGGAGNCT  
 CACTCTGNCACGCCAGACTNGAGTGCAATGGCACCATCTGTGGCTCACTGNNAGCTCTGGTTNC  
 NCAGTTCAAGCCNTTCTCTGCCTCANNCCTCCCAAGATACTAAGACCACAGCCATGTGNTACCGAC  
 ACCTAATAGTTNTNTTATATAGNNAGNTGGGGTTTCANCAATGGTNGGNCAGTACTGGGTCTAAG  
 ACTACT

SEQ ID NO: 501 CCGGCCGNCNGCTCAGTCTTNNNTATTACANCNNTCATTGANTATAAAAAAN  
 TCANTTNNNNTTANCCAATAAAGGNCACCTTCNAAAAGCAAACATTCACTCAGGCNACTAATTCCN  
 CCATATCTCAAACATTTCACTTTTGCCTGATGCCGCAAGCCTGANAGGNGGGTGGCNCCTGTNCCCT  
 ANGGACNGGTCCACATCTAGAACACATGGCTCTATGCTCTCCTTTGGGGCTTANAC

SEQ ID NO: 502 AATAAANACCTTATCCGTGGNCNCGGCCGANGNACATGATNCANATTGGTTT  
 TGCANTTATTAACGANCTGANCTANANATGTNTAAATGCAGCAGANTTATGNCTGTNCTGCANAN  
 ATGCNGTTAGGNNTCNNTATNGTATNTTGATANAAGAACATNTGNTTGANCANATCAGATGAATN  
 ATCTGANTAGGNTNCTATTTACCTTCTGNATTTTAAACGAAAACCCCTNATAT

SEQ ID NO: 503 ACCCGGGGGGCAATTCCTGTCTTNCGGTGTCTNGGCAACAAANCCGTCCAAA  
 CCGACACGCNTGGTATNCTCGCGGTGTCCGGCAAGANACTNCCAAGACANACCCTATTGACTGAN  
 GCTNATGTGAATNCAANGGCCTATNCCCTTNCCTATNCCACTNAACAANAACCTTNTNGACCTTG



TTTAACNAATCATTTTANTATTAACANCTTNTGAAAAGGAACCCATT

SEQ ID NO: 504 ACGCGGGGGGGGTCGGAAGGGGAAAAACAACACTACGGCTGCGGTGTGGTTGGT  
GGTGAGATGACGACCTTAGTGCTGGATAATGGAGCTTACAACGCCAAATCGGTTACAGCCATGA  
AAATGTGTCGGTTATTCCTAATTGTCAGTTCCGGTCAAAAACAGCACGCTTAAAACTTTTACTGC  
CAACCAGATAGATGAAATAAAAAGACCCCTTCTGGACTNTTTTACATCCTCCCTTTTCAAAAAGGGCTA  
CTTGGTGAATTGGGATGTTTCATGAGACNANTTTGGGATTACCTTTTNGGAAAAANAAATTGNATTAA  
GTTTGATTTTACTATACTNTATATTATTTANTCCNCTNGTACCATTACTTTTAACTNNACNTTN  
AATTCCATANNAATTNATGAATTGAAANTTCTATTTNTNAATAAATANCCTTGTTNTTNNCCTTTNT  
TTATATAANTAANATTCCTTGGGGCTTNTATTGCANTTAAGNTNTTTCCGNAGAATAATCCNTCT  
CNCAANTNATTTTGTTCATTGGTTNNNATTGNATNTTCTCCN

SEQ ID NO: 505 GGTACTATATTGTTTAATTACTGGAACATTGTAGTAAGAATTTATATCAGGAG  
TGTAAGTTACTGAATTTTGTATTTTTTTCAGAATTCTTGTTGTTTGTATTTCTTTATATGTTTATATG  
GATTTTAGGAAGAGCTTGTCAATTTATATTAAAGAAAAGCTACATTGGAATTTGCATTGAATCTGT  
AGATAAATTTGTTGGGTATTTTGATCTTCATATTCAGTATTGCTATATTTGAATATTGAGTATATCT  
CAGTTTATTCAGCTCTTTGATTTCTTTCAGCATTTTAAATAATTTTCAAGTGTATAGTATGTGCAGT  
TATTTTGTAAAGTTCAATTTCTAATNNCTTTTTTCTAGGATACTATTAATGAAATTTTAAATTTAATTC  
TTATTTTGGGTATTCCTATAATATAAAAGTCTTNCCTGGCGGCCGCGCGTTCGAAAAGGGCGAAATTC  
AGCCACTGGCNGCCGTTTCTAATGGATTCCGAACCTCNGTNCCCAAGN

SEQ ID NO: 506 ACTACTAAATTAATAAATTTATTCCACTTTTGAATGACAGCCAAAAATCCAC  
CTAATTGATTCTCATTTGGCACGTTCTTCTCAATTTCTGTTCACTAAATTAATAATTTACTAAATTTAAA  
CTGCCAAGAGCCATGATGTTGTCTGCACAAGACAACATTTTCCATCACTTTCAGAAAAGTTATATTT  
GGCATGTTAAGGAAAAAAAATATAATCCCACAGCAGCAGCTATTTAAATAAAGACAGCCACAGG  
ATTCATGCAGAATATTTTAAATATTCCTGTTGAACACAATGATTTAATTGATTTTTTCTTCACGGAT  
GATGCTCCCAAACATCCTATATGCATCCATGGAAATTTAAAGATCCTGGAATAGCGCTTCCATGTG  
GGACATCTTAAAAAGATAGTATTTTGGCTTCAGTGAGTTACACAAATGAATCACCAGTCCTTATTA  
ATTCAATGGGTCTGGTTACAGAGTGNGGNAATTTTCGTCNTTAAATTGCNGAATTTATAAAGAAAG  
AAAATTCCGGCTGNNCCCCCAAAACAAAGACAATTCCTCCATGNAACTGGNTGNATNAAAGNAGACN  
TTTCCCAGGTGGGAGGTGGGGTACCAANGTNTTTTTGTACCTTGCCGGACNCCCTANGGNAATCC  
CCC

SEQ ID NO: 507 CGAGGTACTTTTGTATTTTGATATGGACAGTTTATTCATTTGCATACAGTTATT  
GACTTTTCCCAGCTGATTAAAAGATAGTCAAGAAATTCTGCAATATAGCTGCCAAAATAGACAG  
CTACATTTTATGATATTGTCATCTTTTCTGNTTTTTTTTCTTTTTTTCTTTAGCTATTTTACTTAAG  
CATAATAGCCCAATAGGACATATAAAAGATTATAAATACAGAGCTTTATTATCCTGACGTCTTGGG  
TCTTTTAAGTATATACTTTTCTGAAAGGTATCCATTTTGTAGGCTTGGGTCTTCATGAGCATACGA  
TTGGTTATTTTGTCTGCTGTTCTCAACATCATCATTGCCTGCTGATGTGCCACGATGCTGCTCCAAT  
AAACAGCAATAAGAATGTCTTAATTTGAGCAGTAACATGATTGCAAGAAACCAAGTTTCACAACT  
TGTAAGTCTGTATTTGGGATCTTGGCTTATTTTCCGCCGNGTTTTCTNAGAACTATTCTGGTGAC  
AATTGAATCAGAGTTTTCTTGCTATTGGGGTGGTAACCNTGGAGAACTTTCTAAGGAAAATNAAG  
GAAACTGGACATCTTTGTAAATGGATAANAAAACACTNCNTTTAAT

SEQ ID NO: 508 ACTGGTTGGGGATGGGAATCGTGCTTTTCTTTAACTTCAGTTTACGAGATGC  
TTTGAGAGCGTTAGGCAAAAGCAGAAATAAATATTAGGAGCAACGGGGAAAGCTTTATAAAAAG  
TCATGGTGGCCACTGTTGCAGCTTTGAAGAATGAGTGCTGGCTTGAACAGTTCTTTGCCTGCATCA  
TTGGTAGCTGCACTGAAAGGAAAAAACTTTCACCTTAAGAATTTGAAAAGGAAGAAACCTGGGCT  
CTGGTCTTCATGGCATTAGACTGAGATGCTTAAACAGAACAGAAAGTAATACGCATTTCTGCCAT  
AGGATAGGGAAAATGTAACAAGCTGGTTGCTCTTGAGGTTAGAAAATTGCTGTTTCTCTGTGGATG  
AAGCTGGATTTACTTGAAAATGGGAGAAGTTGGCTTATTGGTTGAATATTGGGACATCAAGCTATC  
TATACCCAGTTTCAGTCGCAACCAGTTTTNCCTTTGTCTGGGGNAAATCNAACCNAANGAATCNTT  
TTGAATCCCGAATCCTAAATCACCTTTTTTTTTTNAATCCCAATCCCANGGGGGCCCGGGGANAA  
CCCCTTGN

SEQ ID NO: 509 ACATCAGACTAGATACAACATGCAGAATGTTTTCTGAACTTATCCGGAAATT  
CCAAAGAAAACATCATGAAACAGCTTACAAAAAATAATATATGCCCTAGTTATTCACCCCT  
GCTTCAACACTGTCAACGTAAAGGCAGAAATAAAGCAAGCTATCAATACCTCAGAACTACTGATA  
TAAGACATCAAATTTCTAAATCAGTGTATTAATAAAGTGAACACTTCCTCTTTCTCTCTCTTCTAC  
ATTTAACTAGAATCATGTTTAAAAAATACTGATATTAATGTGACACTTCAGAGCTACTACTGGA  
AGGAGTAATTCATAACTTCCCTACCCCTCCTTCCATCCCTGCTGATTCAAGAGAAGGGGGAAAAAAC  
CAAGAAAACNAAACGAAAAACCAACCNNGGTCTCTTGNAGAATTGCTGCTATTCCCAAAATGTTG

GCATTGCTGGCATGCCNCAATGTGGGCCCTGGAATANGATTTTGGGGAACCGGCACCGTNGTANT  
CNCCTNTGCGAGTNCCTTGCGGACCCCTANGGGNATTCACCACTGGGGGCGGTTTAANGGANC  
CANCCGGC

SEQ ID NO: 510 ACTTGTAATTAGCACTTGGTGAAAGCTGGAAGGAAGATAAATAACACTAAAC  
TATGCTATTTGATTTTTCTTCTTGAAAGAGTAAGGTTTACCTGTTACATTTTCAAGTTAATTCATGT  
AAAAAATGATAGTGATTTTGATGTAATTTATCTCTTGTGTTGAATCTGTCATTCAAAGGCCAATAATT  
TAAGTTGCTATCAGCTGATATTAGTAGCTTTGCAACCCTGATAGAGTAAATAAATTTTATGGGTGG  
GTGCCAAATACTGCTGTGAATCTATTTGTATAGTATCCATGAATGAATTTATGGAAATAGATATTT  
GTGCAGCTCAATTTATGCAGAGATTAAATGACATCATAATACTGGATGAAAACTTGCATAGAATTC  
TGATTAATAAGTGGGTCTGTTTCACATGTGCAGTTTGAAGTATTTAAATAACCACTCCTTTCACAGT  
TATTTCTTCTCAGCCGTTTCAAGANCTAACATGTGGATTNAAAAATTGNCTCNTTACCAGAATAC  
ATTTAAGGGATGGTCNAAAANTTTTGCNATTGGATAGGCCNAAAAATGGGAACATTTTT

SEQ ID NO: 511 ACATCAGACTAGATACAACATGCNGAATGTTTTCTGAACTTATCCGGAAATT  
CCAAAGAAAACATCATGAAACAGCTTACAAAAAATATATGCCCTAGTTATTCACCT  
GCTTCAACACTGTCAACGTAAAGGCAGAAATAAAGCAAGCTATCAATACCTCAGAACTACTGATA  
TAAGACATCAAATTTCTAAATCAGTGTATTAAAAAGTGAACACTTCCTCTTCTCTCTCTCTAC  
ATTTAACTAGAATCATGTTTAAAAAACTGATATTAAATGTGACACTTCAGAGCTACTACTGGA  
AGGAGTAATTCATAACTTCCCTACCCTCCTTCCATCCCTGCTGATTCAAGAGAAGGGGGAAAAAAC  
CNNNAAAACNNNNCNNAAAACCNCCNNGGCTNTTGGNNAATTGNTGGNTTTCNCNAAAAGGTN  
GNNTTGTNTTGCNNGGCCCNANGNNNGGNCCCNNGAAAAAGNNTTNGGGNAANTNGNNNN

SEQ ID NO: 512 ACGCGGGGGGCCACGTTACGCGGACACGGGAGCAAGATGGCGATTCCGGGC  
AGGCAGTATGGGCTTATTTTGCCAAAGAAAACACAGCAGTTGCACCCTGTTTTGCAAAAACCATC  
AGTGTGTTGGGAATGATTCTGATGATGATGATGAGACCTCTGTGAGTGAAAGCCTTCAGAGGGAAG  
CTGCTAAGAAGCAGGCCATGAAACAGACCAAACTGGAAATCCAGAAGGCCCTTGCAGAAGATGC  
TACTGTGTATGAATATGACAGTATTTATGATGAAATGCAGAAAAAAAGGAGGAAAAATAATCCCA  
AATTGCTTTTGGGGAAGACAGAAAGCCCAAGTATATTCAAACTTGCTAAAAGCAGTTGAGATC  
AGAAAAAAGGAACAGGGAAGAAAGATGGNAAGAAATNCCGAGAGACAGAAATGGGAATNG  
GGGAGTTGGTGATAAAAAACCTTTTGTGCCTTTCNTTTAANAAAAACTNNNAGAAAAACCTGAA  
NAAAAAANAANAAGGGTTNTNCTTGGAACTTTTTTGGTNTNCCCNCCCAANAANT  
TNNGGGGTTTTTGGNCCCTTAATNNCN

SEQ ID NO: 513 ACATCTCTCTATTAACAGGATTTGTTTACACAATTATATTACACTTCACCAAC  
CTTTATACTGCATTTTCAATTAAATACAAAATACATTTACAAAAGAGTCTACCACGGTGTTCCTTCA  
CAATGCCAGCTTAAGGTCTTTTAAACTTCTCTTCTACATATTTATAGTGGTTACATCTTGATTAT  
ATCAACATTATGAGTTTTATGAGTTTATTTTCTAATCAAAGAGAATAGTGTGAGCCTGTTTCTCAA  
CCAAATAGGAAAAACAGCATGTGAGATGATTCCCTGCACATAACCAAGGAATCCTTTTCATGCAC  
ACAACATTGGACTTTTACTTGTGCAGTCACTTAAACATACAAATCATCTTTCCTTGGGATCGTTTTA  
AATTTCTTGAAATACCAAGTGGGTGGAGAGCTTCTTTCAATGAATCCACACATTTAAAATCCAA  
ACGTGCCTCANGCNTCAGGAGNGTATATTTAATTATATTGAGAGNGAGGACGGGGNAAAAAATT  
GCTGAAAAANAATTTCTTTGGCTGACATTTTCAACCTTAAGCTAAAAAGGGGAAN

SEQ ID NO: 514 NACTNNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGAAAACCTAGGGACTAAA  
ATTAATATAAAAAATTGGCATAATGTTGGATTGAATCTACATTTTGGCAGAAAGTTAAACATTCCAC  
ATAATGTCAAAATTATACATCATGCAGTCTGTTTTTTTGTGTTTATTTTGTGTTTTTGTGAGT  
CTGGCTCTGTCAACCCAGGCTGGAGTGCAGTGGCGTGATCTGCAACCTCTGCCCCCGGGTTCAAGC  
GATTCTCCTGCCTCAGCCTCCCGAGTAGCTGAGATTACAGGTGCGCGCCACCACACTTGGCTAATT  
TTTGTATTATTAAGTAAANACGGGGTTTCAACATGTTGGCTAGGCCGCTCTCTTCTGACCTCAGGG  
NGATCAACCCCTCGGCCTCACAAAATGCTGGGATTACAAGCGTGAACCACTTGCCAGCCCAT  
TATACAANTTGNAAAGNACTTTGCACAANCAGTNTTGGCGTGACACCATNTNTACTTGNACC  
TGGTTGAAAAAANGTTTTTTNTTTTATGAATCCGCNTTAAACCTTGGCCCN

SEQ ID NO: 515 ACGCGGGGGGCCACGTTACGCGGACACGGGAGCAAGATGGCGATTCCGGGC  
AGGCAGTATGGGCTTATTTTGCCAAAGAAAACACAGCAGTTGCACCCTGTTTTGCAAAAACCATC  
AGTGTGTTGGGAATGATTCTGATGATGATGATGAGACCTCTGTGAGTGAAAGCCTTCAGAGGGAAG  
CTGCTAAGAAGCAGGCCATGAAACAGACCAAACTGGAAATCCAGAAGGCCCTTGCAGAAGATGC  
TACTGTGTATGAATATGACAGTATTTATGATGAAATGCAGAAAAAAAGGAGGAAAAATAATCCCA  
AATTGCTTTTGGGGAAGACAGAAAGCCCAAGTATATTCAAACTTGCTAAAAGCAGTTGAGATC  
AGAAAAAAGGAACAGGAAAAAAGAATGGAAAAGAAAAATACAGAGAGAACGAGAAATGGAAAAG  
GGGAGTTTGATGATAAAAAANCATTTTGACATCTGCTNTTAAAAAACTGCCAGANAAAACTT

GAGAANAAAAAAGAAAAAAAANGCTTGTTCCTTGAACNTNTTTGGTTTTACCCACCCAAANANT  
NTTNGGTTTT

SEQ ID NO: 516 GGTACTTTGGTCACCACGGGAAACACAACCAACCAAGTGTTTTGCCAAGAGA  
CACAGATCGGAACGCTCATCAAACTGCATGGCTCATTCAAGTGCAGGTTGCCACAGGTCTACA  
GTCACATGGTTAATGAAAATCCAAAACAGGACCAACAGGATGACACATTTTCTGTTCTTAATGCTC  
ATTGCCATTATTGCTAGATCTCACTCTGTAGCCAGGCTGGAGTGTCCGTGGCCAGATCTCGGCTC  
AATGCAGCTTTGACCTCCCCAGCTCAAGTGATCCTCCACCTCNCCTCCCGAGTTGAGTAACTGGG  
ACCACAGGTAAGGGCCACAAGTCTCAGAAGAGTANGAAGTGCCCAGCATNACACAGGAAGCAAG  
TTACAGAACCAANGGCTCTGGTCATAAACTGGTCAACATCTTTTATCACATAAGGGCTCATAGGTC  
CACTTACTGAAAGCNTNAATTTTGTGGGAAGAAANANAAACCNAGGCTTGGGGGACACACNCAA  
ACGGTGTTTTACNCCNCTGAATTTNT

SEQ ID NO: 517 ACTTGGACCATCCACAGCCCAGCAAGGCAGAGCAGGATGCTTCTATTCCTCC  
TGGCACCCATGAGGCCCTGCTTCAGACAGCCCTTTCTCCTCCTCCTCCTCCACCAGGCCTGTGTCT  
CCTCCCCAAAAGGCAAAAAGAGGCACCAAAACCCCAAGCCCAGCCCATCTCTGATGATGAAGCCAG  
TGATGGGGAAGAAACCCAGGTTAGTGCAGCTGATCTGGAAGCCCTCATCAGTGGCCACTACCCCA  
CCTCCCTTGAGGGAGATTGTGAGCCTAGCCCAGCCCTGCTGTCTGGATAAGTGGGTCTGTGCAC  
AGCCCTCAAGCCAGAAGGCGACCAATCAAACTCCATATCACAGAGAAGCTGGAAGTTCTGGCC  
AAAGCCTACAGTGTTTCANGGAGACAAGTGAAGGGCCTGGGCTATGCCAAGGNCATCAATGCCCT  
NAAGAGCTTCATAAACTGTACCTNGTACGCCGGAGTACTGCACCACTNTAAAGATGGCGANC  
CAGAGCCAAAGAAAAATCCTTTGGTTCNNAAAAATTTCTGAAAAAAGGAAGNNTNTNAAGCCTTNA  
AGCNCACAGCNAACAGGCCNTTTGCNAAAAGGGGTATGNGGGAACCAAAAAAN

SEQ ID NO: 518 ACCGCTGAAGACACCCAGAATGAAGGAAAAAAGACAAAAAAGAATAAAACA  
GCTTTTAGTAACGTTGGAAGAAAAATTAGTCAGCGAGTTATTCATTATTTGATGAGAAGGGCAAT  
GATTTGGGAAACATGCACCGAGCAAATGTGATTAGACTTATGGATGAGCGAGACCTGCGACTGGT  
TCAAAGGAACACCAGCACAGAACCTGCAGAGTATCAGCTCATGACAGGATTGCAGATCCTCCAGG  
AGCGGCAGAGGCTGANGGAGATGGAGAAGGCGAACCCCAAACTGGACCAACCTGAGAAAGG  
AAGTATTTTGNCTTCAAATATTGGACAACATGATTTGGACACNAAGACTTAACNGATTTCAGCAGT  
GGATTAAGAAAAAACACCTAGTCCAGATTACNTTAAAGAAAGGNAAAAATGTTGACNTGTCAGA  
AAATGAAATGGAGGAGATATTTTCATNAAATCTTCAGACTTTCTGGAATACTACTTCTCATNTAGG  
CNCAACTGNTCANGNGGGAAANTTTAAGGGGGCTTCGGCCTTGGCCNAATGAGGGAAGGTTTTAG  
GAACTTCGAGACCCGAAGAGACNTTACCAA

SEQ ID NO: 519 ACAAATGAGACAAGGGGAAATATAAATTAGTGGGGCAGGTCCGGGCTCCGG  
TGGGTGGCAAAGCAGAGAACATCTATCCAGTTAGCCTATGAGGCTTGACCCTCAGAGTTGCTGTCC  
TTGGGCTGGACTTGACCAGCTGTAACCTAAGATAAAACAGTCCCACCCTTAAGGTCATCAAATGA  
AAGACACAAGGACAGCATAGCAGAGGCCTAGCTGGCTGTGCAAAAAATGCACGGGAGATTTC  
GTTGGCTAAAACCTCAATATGAGCCACTGCATGGCCAGGCTGTGCAAAGGAGTGGGGCTCCATTA  
ATAGTGCACGGTGTGGCCCCACTGAGGAGGAGAGAGTCTCCGCACACAGCACTGGCTGTACCT  
GACTTTGGTAGACACAGGCATTGGCATGACCAAAGCTGATCTCATAAATAATTTGGGGAACCATT  
GGCAANTCTNGNACCTNGGCCGCGANACGCTTAN

SEQ ID NO: 520 CGAGGTACTTCACTTGGAAGTCTTGTTAAGTTATTGTTTTGCTGGAAATAT  
TCCATCCAAAAATGGCTGCCAAGAAACCTGCCATAATTCTGCATTTGATGGCACATCATACTGTG  
CAAGATAGAGCCAGTATAATGCCAAATTTTGTATCCATTATTAACCCGTAACCTGGGGGCACATGT  
AGCTGTTAAAAATATGCTCACCATCCGGGCACCAAGCAAAATATGTAGAATCAGAAGCCACCGGTT  
TAGAAATAAGTTTGTAGTTTTTCACATCCCACTTNCATTTGCCCTCAGATTTCCAAATCCAGCT  
AATACTAATATATGTCCATGAGGGCTATAGTAGGCTGCATTACGAGGACCAGTTCCAAAGTCAAA  
TACAGGATCACATTTCAAGNTGAAAAATGTCGCTTTGGCAGGCATTAAACCNATNCAGNACAAA  
ACTCAGTAGAACTAGAAATTCCAACCTACTTATAAATGGGGGCATTTTNGNAAATGCACTCAANA  
CTTTCTCATTTGGTGCCATGGAATGNAAGTTGGTCTTCATANTAGGAACTCTGNCTTGGNAAAA  
TTGGNTACTTTTNCACANAGNNTACNTTTTTTCCACATTGGGACT

SEQ ID NO: 521 GGTACATACAAAATCTGAAACTGACACTGTCAGTTCTATACTTTGCACACGTG  
AAGTGTGAGAAATATTTCTTCAGTAGTACTTACAAGGTGACTATATCAGCATTTGGCTGCTTCCACA  
GCTATGCTCAAAGGGTCTTTCCCTTCTTCATCAGTGGCATGTTGATTGGCACCTCGTTTTAGGAATA  
AACATACCTGCCTGTTAAGAACACAGCATTTTTGAGATATTGTGCAAAAAAAAAAAAAAAAAAAAA  
AAAAAAGT

SEQ ID NO: 522 ACCTGCCGGCCCACTGTGCACTGCTTGTGAGAAAATGGATTAAAACTATTCT  
TACATTTGTTTTGCCTTTCTCCTATTTTTTTTTTAACTGTTAAACTCTAAAGCCATAGCTTTTGTT  
TTATATTGAGACATATAATGTGTAAACTTAGTTCCTCAAATAAATCAATCCTGTCTTTCCCATCTTCG  
ATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAAGATCCCAGTTTAT  
GTTCTGTGTCCTTCCATGATTTCCAACTTAATACTATTAGTAACCACAAGTTCAAGGGTCAAA  
GGGACCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATT  
TACATCTTCTACTGTTCAAATAAGAGATTTTTTAAATTCTGAAAACTGCTTACAATTCATGTTTTT  
TAGCCACTCCACAAACCACTAAAATTTAAGTTTTAANCTATCACTCATGTCAATCATATCTATGA  
GACAAATGTCTCCCGATGCTCTTCTGCGNAAATTTAAATTTGTGTACCTCGGCCCGCAACCACCT  
AAGGGCGAATTTCCAGCACCACTGGCGGGCGGTTCTTAGTGATCCCAACTCGGTACCCAAGCCT  
TGGG

SEQ ID NO: 523 ACGCGGGCTATGCGGAGACAAACCTCCATCTCAGCTGTCACCGTGGAGCTG  
AGCAAAGAAACACTGGACACCATGTTAGATGGCCTGGGCCGCATCCGAGACCAACTCTCTGCCGT  
GGCCAGTAAATGATCCAGCCAGCTGCCAGGGCCACTGCCATGACCCAGCTGCTCATGAGTGATAA  
ATGTCTCCCATATGCAGGCTGCCCTTGCACTGCACTGACAACAGGCAGGATGGTGGGGACAG  
CAGGGGGCTACTGCCATCCAGAAGTTACAGTTGGATTGGGAAGAAGCAGCCAGATCCCCGCTGT  
TCTCACTCATCTTCTTTCTTTCTGAAGCTGGAGAGCAGAAGCCCCCATCTTTGAAAAGCTCCTGA  
GTGCAACTTAATTACCACCATGGCAGGGTGAGGGAACATTTGCATCGTCAGCTGCCTCTGGATAN  
ATGTTTTGANAAATTCANGCCCAAANTCATGCNGCNCTATNCATTAAAGTAAAGTTATTTTCAA

SEQ ID NO: 524 ACAGAGTTAACAAGTTTTGAGTTTTTATATAGGAAAAGCCTAGTCAATTCAG  
ATGCTTTCTAGAAAAATTAACATTAATAAACAAATAGAAATCCATGACTAAAGGGGAAAATAAC  
TTTCAAAGTTACCAAATTCGAATCATATCAGAGACCATTATAAATTTCAAACAGTAGATTTACC  
ACACATATTGCATTTTCAAATTTCTAATGTAGCAAAACGTAACCACATAATTTGGCTACAGCTAATC  
GTTTCAGAAAAGTTTAAAAAATTAGCAAAGTTATATCTATAAACTTTTGTAGTTTTCTTTTGCAA  
AGTAAAAAGGCTTAAATCTTAAATAAAGGAAAAACAAACATCCTCTTAAATTTCTTATAAATAG  
CTCTCCAGACATATATTACAAATCTGCTGTAAGCTTTCTTTACCTGAGAGAACTTCCCAGGATCCTT  
TATCCCAAAGGATTACCTTAAAGAGTTCTTCCATCATTTTACTCATGTGAATATGATTAACCTCTA  
TAGAAGTGAGTTGGGACATATGCATTCTTAATCTGCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT  
TTTGCTTAANGAAAAAAACTCTTTGGTAAAAGGCAAAATATTTCAACCTTTTCAAATGAGCTG  
CCTCTGGGGAAAGAGTTGGTTGGGGAAAGAAGAAAGGANAGANGGGAGAGAGAGAGAAAGGT

SEQ ID NO: 525 ACGCGGGAAATTAGAATAAGCTTTTATCAAATAGATAATTGATGCAATTTAG  
GATTCACGCAAGTTTCAAGTGCAATGGCGGTCTTATAGTTTCAATTCTGAAAATAGCAAACCTTAA  
TAAACAGCCACTTTAACTTGTCTGGCAAACAGACCTGCTGTAGATATAGTCTAAGGTAGTTA  
ACCATATAAGCCTTTTCAACTCTTAATGCCCTCCACATGAATCAGCAGTTAAGAAGGTTCTAGAAC  
CCATGAAAGCTTTTGTATGTATTACTAGGTTTTGTTTTCTTATGTTTGTGATTTTACAGTTCTGAC  
TAAAGCTGACCTAAATGGATCAGTTTATGTGTAATATTCTAGTGCTTTAATGACTCTTTTTTCTTT  
GGAGGGAGGGTAACATTATTTGGACAGATGCAGAAGGAACTGTTAGTGAGTCAAGACAAACACA  
TCTGAAATAAAGGAACTGTGTATTAACATGTTAACAATTCATAACTGCACTTTTTATGACATTTG  
AAAATCTATTTATAGGTACCTCGGC

SEQ ID NO: 526 ACACACACGAAGAGAGAGGGAGTTAGCTGAGAAGCAGCATCCAGGGAAAGA  
AAGGTCTCCACAAGCTGAGATGATGGCCCTCTGCCCTCTGAACTGTTCAATTAACAAGTCACCAAGTC  
CCAGTTCTTTGTTCAAATCCTCAAACCTTGAGAGGGCACTGCTCCAATGAGACAGCTACAGCAACCA  
CAGGGAGGAAGTGCTTTGGGCTGCAGGAGAGACAGTCAAAGCCAAATCACGTCCAGGCGGGATG  
AAGTAAACCTGGACCGAGGAGCCTGACCCAGCTGCGGAGCTCACACATGCCCCATAGGGACGGTC  
CTGCCATATACACAAACAGCACACTCAGCATGTGTGCAGAGGCAGGCAACTCCTGAGCTTGGTTC  
TCGCGGGGTGGGGTGAACAATATACTTTAATGGCCAAGAAGATCCCAGCTCTCTGCATGAGACC  
CACATCTCTCCTGGGCTGTCACTCAATCTTGGGAAATACTGGTTCAAGGAAGAGAAACAAGGGCT  
CCATTGTGTGAGAGCAGCAGGCAAGCTCCCCTTCTGCATNTGGACANGACCCCCCTTNTCCTNA  
AANCANGCTTCANCTTAAGTGNCTGACATTGCTTTCTGATGGCAAAGGTCANAACCCACTTTTT  
TGNCATACAAAAAANTGACCCTANCNTTAACCGTAAANAGAAGGTCAAGGGAACACCCCN  
NAATG

SEQ ID NO: 527 ACCTAGTATAAGCCAATCAATCTCATTTTATCATTTAGTTACTGACTGAGATT  
TCAATCTTCTTTAAATGTTCTAAACCTTATCAGGCTCTATCTGCATTGAATAATAAAGCAAATCA  
AATCAAAGTCTTGATTATATTCAGCTAGCATGAACAATGAAAGTGAGTCAAACGTTCTCCTTCAA  
ACTGAGAACTGGATTACATCTTTTGGTATGAGTTCAATTAATTTCAAGTATTCATCACTGGTAGA  
ATTAAATTTCTTCTTTCTCAGGTAAGAATAATCATTAACCTCAAGTGCTTAAAGAATAAAGGCAA  
TAATAGATAAATATATAATTCTATATTTGGATTGCAAGAATTAAAGTAAAGCAAACACATACA

CACACAACCAGAGTCGCCAGTCATTTATCTTCATATAATCAGACAATAATTTTGTATGATAGAATAC  
TTGGATTAAATTACCTTTTTTTTTCTGTGCTTTAAGATAGAAAATTCAAAGCACTAAAATGCCAACT  
AAAGTAAGCACCGGGTATGAAATTGTGAAGTAAAAANATGAATACTGGATATTTTGTAAATAGTT  
TGGGGNCTCATTCAAGAAGAAGGCCTATTTTCAGGAAGGATAATAAATACNATATTGCATAANGAN  
TTTTGGTACCTCNGGCCCGGACCCCG

SEQ ID NO: 528 ACACATTTAACAGTAGCGAATTACACCAAAATGATTTACTTTGAGATTTGAAT  
AATTTGCATAGCAGTAAAATGTGTTTTGTGTAACATACAAATAGAAAAATGACCCAGTATCTTAAT  
TGATACTTACTGGAGAGTATCAGAATTACCCAGCAGCTCTTACAGAATGCCATAAATTCTTTAAGA  
CTAAATATTGAAATCAATTATTTGAAGTAATGTTTCTGATTTACTGTTAAAAAGTTGCTGAGCTCAGT  
TTTTGGAGATATCATTTATGCCTGCCTGTTCCCTTATGACAGTGAGGCCTTCTTTGGCTCCACCTAG  
TATGATAATCATGGGCTCTGTTTTAGTTGATGAGAAGTGGCTCCTATGAATGCCTCTGCTCAATTC  
TTTTTATTTTACTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT  
CCACCTCCCCACAGTGCTGGGATTACAGGCATGAGCCACCACCGCTGGCTCTCTGTTCTTTTCAG  
TGTCTNCGNGCCATCAGTCAGCAGTGCTTACATGTTTAGCATATTTGCATGCAGTTTCTCTTCTGTC  
CCCGANNATATTTTGGCCAAAAAATTGGCAAAAAGTCCTTNGGNCGCGACCACNCCTAAGGGC

SEQ ID NO: 529 AATTGCCCCCTAGCGTGGTCAGCGGCCGAGGTACTGTGTTCCCCAGGCAATTC  
GAATCCAGACTGCATTCTCCTCAGTTTCTCTGAAGCTGACTGTCACATTTTTTAATGCTCTCTGAAG  
AATTTTCTTGAACGAATTTTAAATTGTTTCATATCAAAAAGGTCAACATCTTCACCTGGTCCCTTA  
CTCATCTGAAAAACATNCCAAAATTTCTGGTGCTGATGAAATNGCATATAAATGATGTCTAACAGG  
GCAGCATCACTGATACTTGCACGCTTTCTCACACAGATGGATCAAGTGCTGAACCTACAGATTCC  
TTTCTCTGTGCGAAATTTACAGTCTGCAGGTTGATTTTCAGACAAAAATCCCAGGCCTTCAGGAT  
TGTTGGCAGTTCATTTCATGGGATTTTCAANGATGGTCCCTCTTGATNAAATCAGCAACAGTCCTN  
TATCCACTCTCTTTGGCACCGGCCACATNTNCNTGCNGTCTCACTCTCAAGAAGCTGGNT

SEQ ID NO: 530 ACCATGACATAAAAAGGTTAAAGAACAGGCAACACAATGAGCACTTAAGTTTT  
TAACATGTGGGGAATAGGGCATTTTAAAGGCTGGAACCAAGTTCAGAGGAAACAAGGGTTTGGGTA  
GAGGTANAAAGGTTTAATTAACCTGAAGATCTGCAATGGAGGGAGGGGTGAAGGAGGAATCT  
TAAGACCNANGGAGAAGAGCCAAGGACAAGGTCAGGTTGAATGAGTAGGAGGTGTTACACATC  
CTCTTGCTGTAAGTAGTGGCGTGATCTGATCTGAGAGGCANTAAAGCAGTGATGAAGACAAAG  
TGCAGCNAAACATTGAAGTCCCTGATTGCAGCANTGACTCTGCAACCTTGGATNAATTTTCAA  
TCTTTATTTTGTAAAAACGGAGATAACCGGCCTATAGGGATTTGTTAAAACTGAATATTTTNAAAA  
ANAGCTTTAAATGGTGCCTNACANATGANACCNATTAATACTGTTTATTNGCC

SEQ ID NO: 531 ACTTGGAGAAAGTATAGCAGCAAAACAATGCCTATAGACAACAGGAAACAGA  
ACATATACCCAGAAAAATGCCCTGGCAATCATCAAATCACAGTTTTCCAACATCAATAAAGTGTTT  
AACTCCTCATTTGAAAGATGGTGTTCTTGATTGAATATTGAAGAATTAATAGAGAACTTCAGTC  
TGGAATGGTGGTAAAGGATCAGATTTGTGATGTGAGAATATCTGACATAATGGATGTATATGAAA  
TGAAACTATCCACATTANCTTCCAAAGAAAGCAGGCTACAAGATCTTTTGAAACAAAAGCTCTA  
GCCCTTGACAGGCTGATAGACTGATTGCTCAGCATCGCTGTCAAAGAACTCAAGCTGAAACAGA  
GGCACGGACACTTGCTAGTATGTTGAGAGAAGTTGAGAGAAAAATGAAGAGCTTAGTGTGTTGC  
TTGAAGGCGCACCAAGTTGAATCAAAAAGAGCGCAGAGTGATATTGAGCATCTCTTCAACATAA  
TTAGGAAGTTAGAGTCTGTGGCTTGNANAACATTGAAATACTGAACAAAATCCCTNCTGGGACTTCT  
TCAGAGAAATGGAAAGTNCCTCNGGCCGGGACCNCCTTAANGGGCGGAATTTCCANCANACCTG  
GCGGCCGTTTCTTANTGGAATCCNNCCTCGGTNCCCNANNTTNGGGGNAATCATGGGGCNNTAAN  
CTGNTTC

SEQ ID NO: 532 GGTACCACCGACTTTTCTGTTCCACATAAGCATTTCCCTTTAGGGCTCTAA  
GATGAGGTATCATCGTTTTTAATCCTGAAGAAGGGCTACTGAGTGAGTGCAGATTATTCGGTAAA  
CACTCTTAGGCCTAACCTAGCTAGTCAGTCAAGCAGTAATCTAGCACAACTCTAATGTTGAGATGA  
TGGCCTCGGTGTGAGTGGCACAGCATATAAGGCACATTCAGGAATCAAGACTTTTTTTTTTTTGGC  
CTACTCCTCTCCCTTCTCAAAGACCTTACAGGCAAGGCTGAATTTCTAAAAATAGCCTTATTAGTTAA  
AAACAACACTGGTATAACTAACTCCCATTCTACTTGAAAAAATCTTTGGAATAATGCTTTTTTA  
GATCAAAATAAAAAAATCAAGCTTTTTATAATGATGATAAGGAATTAATTACAAATTTTTAAATTC  
AATATAGTCCATACAAGGCTTATATACTTTGCTCTAAACCTAGCTCACCTGGTCTAGTAGCTACAA  
CATTTANTACTACAGTCAGAAAAATCTAAATTCTAATTGGTAAATTCATGTCCTCAATAAAATTTGGT  
CTGACCAATGAGACAATGAATTA AAAAGACTTTGCAAAGGTTCCAAGCNTCATCCAAATATATA  
CNTGGAATCATTTATAATTTTCATGGGCTCATCCTGNCCATTCTATTCTTAATCCCCCTNCA

SEQ ID NO: 533 ACTTACATGTGTGAACACATATAAAGTGTGAGGTTTACAGACCCTGGCTCAA  
GGACAGTCTAGGATGGGAAAGGAGGTAGGGCGAGAAGAAGCACATATTTTCTCCCTGGTGCTTCA

GCCTCACCCCTATCCAAGGGACAGACATATGGGGTGTGAGAAACCCATCCCCAGGTCCCAGCCTTC  
 AGGACTGGAGTCTTTTGGAGTCTGGTGGAGTCACAGATCCAGTCTTTGGGGGACACTGGGTCTGTC  
 TCCTTTTGAAAGCCCTGGAAGGTGGGAGGTAAGAAAGTAAAGGGAGACAGGTCCCTGCTAGAAG  
 AACTTGACGCCTCGNCCATCGCTGACANTGATGATCTCGGCCTTNTGCTCCTGCTNTAAGGCCTGC  
 AGAGCCCGCAGTAAAGTGGCTTCATCCAGCCCGTGGAACTCCTCATCCTCTGTGTCTTCCCCATTA  
 GTCAAGTTCATACAGGGTAAAGACNGAGTTGTTCTTGGCCACTTCTGGAACNCACTGATAA

SEQ ID NO: 534 ACATACACAATCACTCAACTGGAACAATCAAAACCATCTATGAGTGTGGTTA  
 TTAATAAATAAAATTACGTTTCATACAATGGTAGAAAATGAAATGTTTTTATTAATTTGATTATTAA  
 TACAAAACACACATATATGAATTATATAACCTAGTGTTATATATTTAAAAATCTTTATGCTTGCA  
 ACTGAAATGTCTCTACTCCAAGGGAAGTTTCTGATTTTTAATTTTCTTATTTAAGGAATCTATTAT  
 ATTCACAATGATTAAATGCCTTACACATAGGCAAAAAGCAGACCCAATCCCAGCAAAACAGAAAA  
 ACCATAAGTCTATCATATCACCATATGTTTCACCATATAGTTTTGAAAAATAATCCTATTTGCAGTT  
 TGGTATGTCTTCATATTTATACTTATTATCAAAGTGATTGTCATATTGAGGCACAGAGCTTAAAGAG  
 GAAATATATATTACTTATAGGGGAACCAGACACTGAAACAAGAATATCAATCAATGGCTTCAAAC  
 AAA

SEQ ID NO: 535 GTACGCGGGTGCCGACACCTCCCTCATCTCTCTTATAGTGGAAGGATGGTCAG  
 CATTAGGCTGATGGGGACTGAGAAGGATAGGAAGGGATANAATTGCCATGTGTATAAAGCTTTA  
 TTCTTTAGCCCTTAACCCCTAAGGCTCAGGGAAATACCCCTATGTTATTGTGCTCCCTGGATTCTGCA  
 ACTCATTTTTCTTCCACTCTGGAGCAGGGTGAGGGGAATGTTATGGGTAAACAGACATGCAGGCAT  
 GGTCTACCCATTCTTTGCAACAAGTATGGGGCCCATGTGGTAGTCCCCATACCCCTCCAGTTCCCT  
 ATATTTTTGTCTTCTTCCCTTTCCCTCTTTGCCATTCTACCTTGCAATTTTTCTGTGCTAGTGCCTTAGC  
 CAAGGCAAGGAGATAAGGATGCTCTTCTTGCTTTTTATATCTGCACATTCATACCTCTCCAAAGAC  
 CAGCTTTTCCCCAGCCAGGGCCCTCAGCCTTCTGCTGCCCCAGTGATTGATTGAGAGAGCTGTTG  
 GGGTTTCTTCTGCAATGACCCCTGGGAAGANGGACCTTTGGTAAGGGNCATGATAAANTGGCGG  
 GGGTCTTGGTCTGCTCANGNTTTTCATNCTTCTCTCTTCTCTNTGNGACTGNGGATATGGGTA  
 TTAAGGGGGGTGCACCCNGGGANCCCTGANAACTGGGTGCNCAAAATTCAAAAANAANGNG

SEQ ID NO: 536 ACCTCTACGTATTGACAACCTTTGAGTTCTGTAGATAACAAGCAGATTTGGGTC  
 TCCTGTGATTGGCTAATGGTCTCCATCTCCCAGCAGACTTAATTCAGGTTTTGCTTCTGCTACATCC  
 CGCCAGTAAGGAAGCAGCAAAGGTAGAGAAGAGACCTTTTTCTCTATCAAAGGCCAGAGATGCGA  
 GAACAAAAATTCAATCCCCTTTGGAGACAAATGTAGTCCATCTGATAAATAAGATGAGAAGTCCT  
 GGCTGCTCTGCATCAGGGTCCACAGGTCAAGTACC

SEQ ID NO: 537 GGNACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACNCNGAAAAATTG  
 GAACATANTTNTNTTGGCTGGAGGTTTGGACATTCTANAGCAATACATATGCCTTTCAACTNTTG  
 AAANACCTCACTACCGCCTCCTTNTTGAGCTTTTTTGGAGGAGCATTACACANAGCATTNTGGC  
 AGCTTNTANTTNTGANATGAGGTATGTGAGCAANAGGAGGCNNTTNTNTGAATGAAAAGGCGCT  
 TGGTCACATCCCCANATGTCAAGTGAANATACNNGCAGTTCATTTCCCTAGTAGCCCACTCACC  
 TCAAGCTGGAATNTTTCAGCTTCACTCGGACTGTTAGTTGCTTGCAACGTTTTCTCTAGTTTACAN  
 AGCACTNTTAATTCAAACNCCANCCAAGCACAGAGTTTGGTAAACTCGGGGGAAGTGGCTCCAGC  
 AAAGACTGCCTGAGANAGCGCTCCNTCTTCCAACAATGGGCCCTTGTAACTAAATNTTCCAACG  
 ACTNCAAAATGTCANTCTNCATGAGGGNCACACTCCATGCTACTGNGGGTATTCAAATTTCCAGN  
 CGNCAGGCTCCCCCTGTTGCGCGGAACGCNTACCTGNCCCGGGCGGGCGGTTTNAANGGGCGN  
 AATTCCANCCCCCTTGGNNGGCGGTTNCTTAANGGGATTTCNNAGCCNNGGTNCCAANCTTTG

SEQ ID NO: 538 ACTGGGATTACAGGCGTGAACCACCGTGCCCGGCTTCCCCAGATATCTTCA  
 AAGCAACTGCTAGTCTGCTTTTGACATTACACTCTACATTCTCATTGCCCACTTAACCTAACCCCT  
 CCCTCATGTCTTTTCTAAAGATGTAGAAGACTTTGAAATAGCATAACATTTAAATGGTCACAAAT  
 CACCTGGCTTGAAAGTGACTAATAATAGGAAAATCAGACCTCAGGACTGAGTTCAATTTCAAGAT  
 TCTCCCACTCCACTTTGCTAAGAGGGAAACAGAAAAACATTACTCCAGCCACACAAATATGATAC  
 TGTGCAAGTATCCCACTCTTACTCATTCAAATCCTACCCATTTTTTCCCATCGTGGAACAGAGTCTT  
 GCCAATTATACCTCTCTCCAATAACCCATAAAGATGGTGATTGAAATTTGGGCCAAGATTTAACA  
 TTTTATGGAATACAAGAGGAGAGTTTCTGCAAAACACAGGATTGTCAGACAAGGAGACTTAGGT  
 TCAAATTCAGTTCTGCCATTTATTACCTCAAANGGGAT

SEQ ID NO: 539 GGTACGCGGGGAGTTTAGCGTAAACCGGGAAGCGGATCGCGTGAGTGAAA  
 GTCACCGCAGCGGAGATGGACAAACCATGTGGGTGCCCTCCAGGTGTGTGCGACCATGAAATGGG  
 AGACTAGCAGGACCCACGTTGGCCAACCGTGGACCTGGTCCCTCCAGTAGGAGTCGTGAGCCAGC  
 TGAATTTGAATGCAAAGATGGAGCAAGGGCCGACTGGAGTCACACTGACATCCACCCCCATAACA  
 TGGGGACAGATCAAGAAAACAAAGCAGGAAGCTGAGAAAATGCTGGAGCGTCCATGTGCTTGGC

CATGTTAGCAGTAGTGTCTGTGAGGATTGATTTCCAAAACTCCTGTTACATGAGGAAGACACCA  
AGAAGAGCAAAGAAAAGGAGGCCAGGGTGGCTCTTCTCAGATGGAGTCTCGGTCTGTCGCCCAGC  
CTGGAGCGTAGTGATGCGATCTTGGCTCACTGCAACCTCCACCTCCTGAGGTGGGAGAATCGCTTG  
AACCCAAGGAGGTGGAAGTTTGCAGTTGAACCCGATAGTGCCATTGCACTCCANCCCTGGGCAAC  
AAGAAGTGAAAACCTTTTTCTNAAAAAAAAAAAAATA

SEQ ID NO: 540 GGTACAAAGTCCAGAATAACTTGTATCACCCTGTGTAATCAACAAGCTCTT  
GTCTCATGGAGTGCCTCTGCCTAATTGGCTTATAAACAGTTACAAGAAGGTTGATGCTGCTGAATT  
GCTTCGTTTATACTTAACTGTGACCTTTTAGAAGAAGCTGTGGATTTGGTGTGAGAATATGTGGA  
TGCTGTATTGGGAAAAGGACATCAATACTTCGGAATTGAGTTTCCACTGTCCGCAACAGCCCCAAT  
GGTGTGGCTTCCATACTCCTCTATTGATCAGCTTCTCCAAGCTCTGGGAGAGAACAGTGCCAACAG  
TCACAACATCGCACTGTCCAGAAAATACTTGACAAAATTGGAGGACTACCAGCAAAAAGTTGATA  
AGGCAACACGGGATTTATTATATCGTCGGACCTTGTGATTTGGATTGTACCTAGCCTTTGTAACC  
GCTTGGTGCCTCTTAGGACTTAAGACTACCCTACAGGAACCCTGTCTGCCCGGGCNGGCGCTCG  
A

SEQ ID NO: 541 GGTACATGAGGCTTCCAGCCCAGCCCTAGGAGATCCATCCCAAAGACCCAC  
AGAGACCTGCATGGGAGGTGGGGCCACAGGTCTGGTATCAGGCAAACCTAGGTTGGAACACTGGC  
TCCATAAAGAGGAAGTCACTTAACCTTCTCTGGGGCATGGTTTCTTCATCTGTTCCCACCTCTGAA  
GACTATCGTAGACAGAATGAAAGTTAAGCAACTTAACGCAACGCCAGGATACCAGAATTATTC  
TAAATGGCAGAACTCCTACTTAGTCTGTATCTTGGGAGTTCTCTAGGCAGGCAGGTTGCCAGGGT  
GGGGCTGAGATCCAGATGTGCTCTAGGTCCCTGTCTGCTGCAGAATCATGTGGCTGCTGGACCTGG  
GGGTCCCTCANGTCCTTGCAGGAGCTGAGGGTAGGAGACTCCATTTGCCAAACAACCTTANAACCTT  
GGGCCTCAGTTTCTTGTCTTTAAAAGGAACAGGTAAAGAATTCATAAAAACAATTGACAANGGG  
CTGCCTGGAATTTNCCTGTAATCTCTTCCAGCANAAAGANGCCTCANANGTGANAGTAAGAACAGG  
AAAACANGGGACAGGTTTCACTCTGTGCCAAGTGTGGTGGCTTAAGGTCTGTAATCCCANCNCTTT  
GGGAGGGCTNAAATGGGAGGAATTGATTGATNCCCAGAAAGTTCAAAGGCTTCNATTAAACCTGTG  
CCTGGGGG

SEQ ID NO: 542 ACAAAGACTTTGTAAATGTGATTCAGGGCCCCCAGCACCCCTGTGTCTGCAG  
AGTGCCTTCAAAACTCAGCTGTTCCAGCCGGTGCCAACCTGTGAACCTCCCACCATATCCCAGAAT  
CTGCTATTCCCCAAACCACTTCCCAGTTTCCTTTAGTAATCTTTCTGAAGGAGCCAGGACAATAG  
GGCCTGTTGTTTAGTGAATTTCTTTATTATTTTTCAGCCTTTAAAATGTAATTTCCATCTCTTGCAATG  
AATTTGTTTCCCTTTTTTTTTGCTTCATTTGTTTAAATTTTTCAGGTATTTAGCTCCCCTTTTCATATTAT  
TTTTAAATTTTAAATTACCTGTTGTAGGGTGTCTCCAGAAGCAAAGAGCAAAATTTTACTGTTG  
TGATGTACC

SEQ ID NO: 543 GGTACGCGGAGGAGGGCCCATGTGCTGAAAATCCGAAGTGCCGCGGAAAGT  
GGAGAGCTGACAAGGAAGGTTTCGAGCGTTTTGCTGGCAAAGGGATTTCTTACAACCTCCAGGCA  
TGCGTCTTCTGCCCTGCTGGCCTTGGCATCCAAGGTCACTCTGCCCCCATTACCGCTATGGGAT  
GAGCCCCCAGGCTCTGTTGCAGACAAGAGGAAGAACCCCCCATGGATCAGGCGGCGCCAGTG  
TTGTGGAACCCATCTCTGATGAAGACTGGTATCTGTTCTGTGGGGACACGGTGGAGATCCTAGAAG  
GCAAGGATGCCGGAAGCAGGGCAAAGTGGTTCAAGTTATCCGGCAGCGAAACTGGGTGGTGTG  
GGGAGGGCTGAACACACATTACCGCTACATTGGCAAGACCATTGATTACCGGGGAACCATGATCC  
CTAGTGAAGCCCCCTTGCTCCACCCGCCAGGTCAAACCTGTGGATCCTATGGACAGGAAACCCACT  
GAGATCGAGTTGGAATAATTACTGAAAGCAAGGAGANCCGGGTACCTGCCCGGGCCGGCCGCTCG  
AAAGGGCCNAATTTCCAACAACACCTGGCCGGNCGTTTACTAGTNGGANTCCGNAGCTCCGGT

SEQ ID NO: 544 GGTACTTGTTTAAACCCAGAGTTAACTACCCTGGATAGCACAAATTGTTTTGGA  
CTCAGAAGTAAATCTGAAGCCTGTATGTTGGTTTACTGCTAAATTCTGATAGTGCTATTCTAAAAG  
ATATTACTGCTGATGCAAAGATATTTTTCTTAGAGGAATTATGAAGGAAGAGTGAAAGAGGATGA  
GGGTGTAGGGAGAAGGAACAAAGATTAGAAAGAGGGAATATTTAGGTTCTTGATTAAGAGGCAG  
AGCTGATTTAAAGATACCAATATGGAACGAGTGCTGGGTATCCACATCTATCTTTCTTAATCCT  
AAGAATAAAGCCATTTCAATTAGATCCAAATAGTTAATAAGAAATCTGCTAATTATGGATTCTTTT  
TTCTTTTTTTTTGCTACCAAACCTACAGATAACCATTTCTAGAAAAGAATAAGACTGATCCATTTCTG  
CCTCTCATGTGGAGCCTAGGATTCTTCANATAGTAAAAGACAGNTGGAGTCTGTGAGGACAGCTA  
CCGCCACATTTTGGCAGCATTTCTTACCAGATTGCTGATTNCATGACTTGGATTTTTTTTTTCTTTT  
TCTTGGAGGCATAATTTCCGAAAGAGTAAGATAAACTTTCTTATTAATAAACTGGGTTTAGGTCCAA  
ATTATGGAAGATGTTGAAAAACANCTCCAGTCCCCTTTACATTTTGAANTCTTTTTAAAAAATT  
TCA



SEQ ID NO: 545 ACTCACAAGTAAGAACTTTCTCTACTGAAGGATACTGTACAGAGTTTGTG  
CAGAGCATCTATATATATATTTATTTATTTATTTTAAAAAATAAACAACAATGATGAACGAACCC  
AGGTTCCCTAGAACCAATTCTCTTGATTCTCTACTTCCACAAAATAAAGTGTATCATTGGCCAAGA  
CTACAGATGTGTTTTTTTTTTTT

SEQ ID NO: 546 GGTACGCGGGGACCTCGGTGATGTCGTGGGTTCAAGCAGCCTCCTTGATCCA  
GGGCCCTGGAGACAAAGGGGACGTGTTTGACGAAGAAGCAGACGAGTCGCTCCTGGCGCAGCGG  
GAATGGCAGAGTAACATGCAAAGACGAGTCAAAGAAGGTTATAGAGATGGAATAGATGCTGGCA  
AAGCAGTTACTCTTCAACAGGGCTTCAATCAAGGTTATAAGAAAGGTGCAGAAAGTCATTTTAAAC  
TATGGACGACTCCGAGGAACATTGAGTGCTTTGCTCTCCTGGTGTACCTTCATAATAATAATTCA  
ACTTTGATCAATAAAATAAACAATCTTCTGGATGCAGTTGGCCAGTGTGAAGAGTATGTGCTCAAA  
CATCTGAAATCAATCACTCCACCGTCCCATGTTGTAGATTATTGGACTCCATTGAGGATATGGAC  
CTTTGTGATGTAGTTCCAGCTGAGAAAAAGATTGATGAAGCTAAAGATGAAAGACTCTGTGAAAA  
TAATGCTTGAGTTTAACAAAAACTGTAGCAAGAGCCATAATGGGGATAGATTGGTCATATGTANA  
AATGTTGTAGAACAAGGAGCATGCACATTGAGAAAACCCAAGCCCCCATGGAATTTTTGGAA  
CAGACAGCCAGTTTAAATTAACCAGCTGGGCCNTTCA

SEQ ID NO: 547 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCGAGTTTTTCATTTTTTTATTTA  
ANAAAATACTTGAATTGCCTTAAACAATNTTAAATNTTTAAACAACATGANAAANAGTGCCANAG  
GTCAAAACATAGTATTTAGTTCACTGAGTTGCCCTGACANATAATGAATGGGGATTGATTTAATAG  
TGACCAAATACACTGGCCATNTTACTAAAGNGCTGTAAAATGGCCAAGNGAGGACAACCTGCATN  
TAAAATGANATCAAATCCTCGAGTCCATTCTTTTAGCAAAAATGATTAAAACCATNTTGGCAGGA  
CCAAGTNTTTGCAAAGCCTATCAAATGGATGGATGGCTTCAANACAGCANANAACCCACAANCTT  
GAAAAGGCCTCCGGANAGTTCCTCAGGATAAACGGGGTGCTGGCATCGTCTGGTCTCT

SEQ ID NO: 548 TGTACTTNTTTTTTTTTTTTTTTTTTTTTTNCATTGGTTAAACAGTTTANTTCCC  
AAAGCTAGTAATTTTAGTTAAATATNCATTANANCCTTTTTANATGGCTGCTAATAAACNCTNTGT  
CAAAATGNGTAGTTTTAAACTCAAACCTCGAAAGCCAANATAANCAACTCCTTCANTTNTTACTCTG  
ACCAAGGCNTAANAATTCATTANACAAAAAGCTTTCAAAACCTACCTAAAAATAANATAGTTCA  
TAAATTTTCAAACTGTNTTCCCTGTTGCGGACAGCCCTTGATCTTTGTAAAACTTANCAAAATTT  
GGCATGCTCTCATGTTAGCTTTTTAAGTTACTGAAACTATAAATTTAGCNTCATTNTCAAATCTG  
TATAGTTTCTCATTCGGAANGCTTAAACATTTAGGTCAAAAATTAATAATCCAGTANAAAAATATT  
CNTCTTTAGCCAGGTTGTATAGTTAACANCATGGCAANAACTGGTGAGAACATTTAANCATTCA  
GAGCATAAAAAATACTTCAAAGCCTTNCGAAACTTGAACCTTAAGCCATTTTCTAAATTTTNAAGTAG  
TTCCAACTACAGGGAAAATAAAAACTGCCCTAGGGACTGGAGAATGATTTAANCCNCCTTCAAA  
ATTTTTTCCNGGTTCCCTGANATNAATTTTNTAGGNTTNTTNTTGGNAACCAAAAAAAGNTGGTTA

SEQ ID NO: 549 GGACCCTTATATTTCACTACTTTCTGTTTCATAAGTTATAGTGACATTGCTCTTT  
GGTAAAAATGCCTGCTTTCCAATACTTTGATTGCATATTAGACATTCTTAACAGGGCGGCAGTCTA  
GTGTGAAAGTTTTATTTTCCATTTTCTTTAAGTAAATTTTTTAAAAAATCTGATTTANGGC  
TAGGTGTGGTGCTCAGGCCTGTAATCCTGGCACTTTGGGAGGCCAAGGTGGGAAGATCGTGTGA  
GGCCAAGAGTTCAAGACCAGCCTGGGCAACATAGCGAGACCCCTATCTGTATTAAAAAAAATCT  
GATTTAATCTTTTTATTTATCATAAGGGGTTAATTCCTGAAGTAAAGGTTTGCACCTATTAACTT  
AAAAGTCCAAATGATTTTGTTCNTTATGTGCGTGATAAAAATACAAAGAATGGTGTGGCCACC  
TCCTCCCTTTCAAGCTNGGGCAACAGGTAGCTCTTTCCAGCCCCTGANCCCAGCCCCTTCCCAAGT  
GGTGCCCGGACAAAAAACTACATGGGCCNTTTCGTGTCTTGGGGGTGAAAAAGGGAGGGATGAAT  
TGGGGTGANAGAACCCTGTNGAATT

SEQ ID NO: 550 GGTACCAAATTTGCATATTTGAAATTAACACTTTAGCATTTGCTGAACTCAGC  
CCTCGTTAACTCCCTTAACAAGTTCAATCTGAAATCGAATTTGCATTCAAACAGTTTAAATGCCACC  
AAGTAGGTCTGAACTAATGTATAAACTCAGCGCCGCCGCCGCCACCCCTACTTTCAGGGCAGCTG  
CTCGGGGAAGCCGGTTTTTTTTTTTGGCCATTTGCAAACAAAACCAACCCACACCCGTTATCGCAG  
AGCACCCAAGGCCCATGGCAACTTGTTCCACAAGGGAGAGCCTTCCAAGGCCATATTGTCAGTC  
TAATTAATATGAGCTTTTTTTTTTTTTTTCAGTGCTGCCTATGTTTCTTGTGCTGGTANAATGTGGGG  
GCATTTATTAATGGTTGCTGCCNAAAAAAAANGAAAGGAAANGAAAAANGAAAAATCCGAAAC  
ACCCCTCCCCNGAACCACCCCAATACTGCTGCGTGGAATGAATCGGCATTGTTCTTAGAGTTTG  
GCNCTCTTTTTTTTTTCTGCATTCAATCTCTTTTTTGGCAGGACNTCACNTTGTNTGANNAGGGCCTT  
TATTNTGTAGAGTGCACCATTTGGGGAACCTTTTTTGAAAAAANACCCAAACCGGGAANGTTTTTT  
TGATTTTAAAAAAAACNNGGGGGATGNTCTNGAAAGGAAAATGGAAATGGAGNG

SEQ ID NO: 551 ACCCAAGGGAGGAGCCNAACCTAANCGGCGGAAGAAAGTGAGGAGGCCCTT  
CCAACGTTGATGCCCCCTCTCTTTCCTCAAATCAATGTCAGGGAGTCAAAAGGGCTGTAGCACAGG



ATGGAGTTTGATTTATNCCTNCTCCCCAACACCTAGGAACTGAATCTTTTTCTTTTTATTTTTTGA  
GATGGAGTCTTGCTCTGTTGCCAGCTGGANTGCAGTGGTGTGATCTCANCTCACTGCAACCTCTG  
TCTGCC

SEQ ID NO: 552 ACATGGCTCCATGGAGGTTCTCCAGTCGGTGTGCTGCTGCTGTTTTTCGAGCC  
TTATCTCGTCTGTGCTCCTCATAAGTGTCTCATCAGTGGGCAGCTCATAGCGGCACAAGGGACAG  
GAATTTGTCTTGCTTAGCCAGGGCAGAATGCAGCTGGAATGGAAAAGGTGATGGCAAGGCATCTC  
AATGGCAGTCTCCTCCTCCTCAAATTCAAAAGACACACGGGGCACTTGAGCTCAGCCTGAGAGC  
CTCTGATGACTGTCTGGGGAGGTTCTCAACCACAGTCTTGGCAGCTGGTGGAGGCAGGTGGTGGT  
CCCAATCTACTACCAACCCCAAGTCTTCAAAGTCCATCCTATTGAAAAGTGACCTTGCGAGCTCCA  
GCAGCATGTTGGTTCGCGTCTCCTGCTCANGGTCCGACGGCTCGCAGTCCGTGTTTCGAAATAG  
GACGCCATGGCTGCCACCCCTTCTGACACAACCCCCCNGCGTACCTNNGCCGCNANCACGCTAAN  
GGGCGAAATTNCANCNACACTGGCGGGCCGTTACTANNTGGGATCCCGAGCTTCGGGTACCAANN  
TTTGGCGGTAATCATGGGCATAAACTTGTTTTCTGGTGTGAAAATTGNTATNCGCTTCACAATTC  
CCCACAACATNCCAGCCCGNAAGCNTTAAAGTGGTAAAAGCCCNGGGGTGCCCTAAAG

SEQ ID NO: 553 ACAAATTGGGTCAAATGGCTGCTTCTCATGGCTCTTGGCTCCAAAAGTTACAG  
TTCCAGAAAGACTAATTTCCATTGATTTGGGAACCTTCTGGCCAATAATCCAGATCAATAAGCTTT  
TCTCTCGAAATACTTCAAGCTGGCCAAAAGTATTTGTATTCCAAATGTGTAATTGGACCTCTATT  
GTAAAAGGTATATGGGCTTCACAGAATTCAAAATTATTTTCACACTTTC

SEQ ID NO: 554 ACATGATGAAGCACTTACAATAGTGTCTGGCACATACAAATACTCTGCAAAT  
ATTGCTTATTATCATAAGTCTCTAATATTCTGACATCTGAAGCCTTTGGGGTTCTAAATTAGCTGTT  
TTTTATGCTGCCACTTAAGGTGTTTTTTATTTTTCAACGTCTCTGATCTTTGATTGTGAGCTCATAT  
TTGGAGGATCTTTGGGAATCCTGAGGCTCTAAATTTGGGAAATTTTCTCCAGACCATCTGCAATT  
TGAGACCACATTTACTGCCATCAGTTATCCCTGCTTCATGTAGGAGCCCTGGTTCAGCTCCCTTGTC  
TTCAGTGTGTCCAGGTTTAGCTCCACTGCCAGCCTTCATATAGGTCTCGACCCAAGATTTTAGTGTA  
TTGCTCCCAGATCATCACCTTAGTTTCCACTCACAGTTCTGATGGTTTTGTTTTATTTTGTGTTT  
TCTGAGTTTCTATTTTCTTTGTTTTTGGCATTGTGAGGAATCTGTTGTTTCCTAAGAACCTANTACCT  
CG

SEQ ID NO: 555 ACGCGGGGAGCGCAGCGGGGGCGGGAAGGTTGTAGTGCCGCGAGTTGAGCT  
CCTCTTGCTAAGTGGTCGCGCCCCCTTTAAGAGCAGCGATTGTAAGGAGAGGGCGTCCCGGTGTC  
CTCGGGTCCCAGGTGATTGTGAAGTGCTGACCAATTGCCACTGGACATACTTGAAACAAAATAGG  
AAAATGGCAGCAAACTCTTCAGGACAAGGTTTTCAAAAACAAAATAGAGTTGCAATCTTGGCAGA  
ACTGGACAAAAGAGAAAAGAAAAGTACTTATGCAGAACAGTCTTCAACAAATCATCTGGAGCTA  
GCATTCGACTCTCGAGACCCTCTCTTAATAAGGACTTCCGGGATCACGCTGAGCANCAGCATATTG  
CANCCAAACANAAAGCAGCTTTCNAGCATGCTCATGCACATTCATCTGGATACTTTCATCTCAA  
ACTCTTGCAATTTGGGAACCTTATTTTCTGTTTTACCTCGCCTTTACCCANAATGAAAAAAACAT  
TTGCGATNGAAAAGTGACTTTNTAATATCAAAATGCCNAAAGCTTCTATCATTCATGCTTCATT  
GAACTGGTGACTTTTAAAAAATTTT

SEQ ID NO: 556 ACGCGGGGGTTTGGATCCGGGTGAGTCCGGGTGCGGAGATTGAGGAGAGACGC  
TCTGAACTGACTGCCCCGCATACCGGAGCGTCCCAGCTGCGAGGAGTGTAACAGGAACATCGA  
TAAGTAGTGTAACAACTTGACAAATGAAATCCGAAGCCAAGGATGGAGAGGAGGAGAGTCTACA  
GACTGCTTTCAAAAAATTAAGAGTGGATGCATCAGGGTCTGTAGCATCTCTGTCTGTTGGAGAAGG  
CACAGGTGTCAGAGCACCAAGTCAGAACAGCAACAGATGATACCAAACCTAAAACCATGTGCAT  
CTAAAGACAGTTGGCACGGGTCTACAAGGAAGTCTTACGAGGAGCAGTGAGAACTCANCCTCGT  
CGACGTTCTAAGTCTCCTGTCCTTCATCCTCCAAAGTTTATACATTGCAGTACC

SEQ ID NO: 557 GGTACATAAAGATATGTAAACCACATTAATCTTGCAGTAATATAATGTCAAC  
ATCATCATGAACACCTTGCGAGAAGCAGCTCTCCTGGATATGTTACAATCTTCTCGTTTTGCAGCT  
TTAAGAGTTTCTACCAATGCTTCAAAGAGGTTGGCACATTTATCATCACGGAAGAGGACCCAAA  
TTTCACGCTTAACCTTTCCATCAGCATTTTTTGAACCCAAACGATGAATTTCTTCCACTAAGAGGTTA  
ACCTCGTGATCCACATTCATTGCTGCCATGCGTAGGTCCTCTTCATGGTATCCTGCCAGGCACCAA  
AAATACAAAAGGTGTGTAGCACAGGATCTGCCCTCGGGGATCTCACATCTAAAGGCAGATGAGG  
GCCTCTCTGCGCGGTTTTCTTCCCGCCGGAGAACTGGCGCCGCGTGGTCTGGACCCCGTTTACCT  
TCTTCAATGCACTGGGGAGGGTGAGGCGAAGGAAGGAAAGTGGCTGAGGCACAAAANAAAAAACN  
GCTGTGGGAGGAAGAAGACCAGGGTANGGGGAAAAAGTAATGTTTCTTGCCTCAAAGAAAGGG  
GCCGGCCAAACTTCCCCCGNGTACCTTGCCCGGGCCGGGCGGTNCGAAAAG

SEQ ID NO: 558 GNGTACTGGTTGTTGTGGAGGAACCTCTGGCTTGCTCATTAAAGTCCTACTGAT  
TTTCACTATCCCCTGAATTTCCCACTTATTTTGTCTTTCACTATCGCAGGCCTTAGAAGAGGTCT  
ACCTGCCTCCAGTCTTACCTAGTCCAGTCTATCCCCTGGAGTTAGAATGGCCATCCTGAAGTGAAA  
AGTAATGTACATTACTCCCTTCAGTGATTTCTTGTAGAAGTGCCAATCCCTGAATGCCACCAAGA  
TCTTAATCTTCACATCTTTAATCTTATCTCTTTGACTCCTCTTTACACCGGAAAAACGGCTCCAGCTG  
TTCTAGCTCTCTTTTCAGTTCTTTGAACCTTCCCACTTAGGGTCTATAAGGTTCCCTCTGCCCAAAT  
TGTTCTACTCTCCCTCTTCTTCAACACATCCTTCAGTTTAAAGCACTTGCTTCTCTCAATTTAAACT  
CCACTTCCTTAGGGATGTCTCTGTGACCTCCCTGTGCTGGATTAGCTTCCTACTCTATGCTCTTGTA  
AAANAACCTATCTACTTCCTTTGNCATAAACTCATCATTTTCTATAAACATTAATATTGGCCTCCTG  
CTAGAATGNAAAGCTCCCGTGAGAGCAANGATCCCTCCTGTTTACCCTGT

SEQ ID NO: 559 GGTACACACACACACTCTCTCTCTCTCTCACCTGGCTAAGGCTTTTAAA  
ATTATGAAGATAAATAATCTGTTTCACCAGCTGGAGTGAGTTGCAAGGAAGATTGCTGGAGCTAC  
TCAATCTAGCGCTAATGGTTTGGATTCTTACTGCAAACCTACATAATTTAACATATTTGTTTACTT  
TAGTGTGACAACTGATGAAAAAAAATGGAGCAATCTGAATTGTATAAAAATAACTTAAGAAGGAAG  
AAAAGTGATATATAAATATATTTTGCAAATGTCACATTAATTTAAAAAATGAGTATGATTGATTTTA  
TTTTTAAAGTGGGCATTCTTCACTGTTTCGAGACCTTTGTATGTATTTGTGTATTTTTTATCTTTTTT  
TTCAGGCCATTATTATAAGGTGTTATTTGGCCCTCNAATGTAGAAGTTATGTTTAAATCAACCAA  
TCAAGGCCCT

SEQ ID NO: 560 ACTTTTAATGGTGGGAATTTACAGTAGAAGCATCCTTTGCTGAGTTATACATT  
CCTTTATCAATCTCTTTTGATACAACATTTAAACAAGTAGCTTCAAGAAACCACTGGTGTTTGA  
GGATAGTATTTCTAAATAGCATTCAGGAACAGAGTATTATTGCACAGATCTGAAGATCAAAAAAA  
AGCTCAAGGAAATACAGATCGGAAGTGCTGATGAGTTATATTTATTGAAAACCCAACCTTTTAAAG  
AAGTGCTAAGATCAGTCACCCATGTGAATAAGAAGCCAGGAAAGGAAAGATGGGGAAGCCCAGA  
TCACCAGGCTTCTATTAAGGAGGAAAGCAACAGAGGAAACAGTGAAGGGGAACAGAAGGGGGTA  
GCAAAGTGTTACAGAAAAGCGGACTGGATAGACAAAACCTGCAGAAGGTGTATGTTGGGGAGAAC  
TGAAAGGGGAAAACAAAAATACTTGACATAGTCTTAAGTAGAAGAAGGCNAGTTAGAGAAAAACA  
AAAGTATCTACTGGCCTTGTCAACATACAGACTTCAAAATCCCCTTATGAGAATCCAAAGAATGAT  
GTGTGTAAGGGAAAATTTATTTGCCCTTCCGGGAAGAAATCAGTATCTTTGCCAAATCTTGGA  
GACGAAATCAAAGCCCCATTAATGATTCANAATCAGTGGCTTGACCTCCTG

SEQ ID NO: 561 ACTTTNTTTTTTTTTTTTTTTTTTTTTCTAAAAAATAAAAACTTTATTTTTTTTCAA  
GTTTATAANATAGTTCCCATTACATATAACATTACGGTCACGGATTCTACAGCCACAAATGCCCGC  
AGTCACATAAATATATCCAATCCAATCAATGCCTTTTCTGCTAACANAGGCATCTGAAGTTCAA  
GGGANAGTCNCATTTTGAAGTAAAGTCGTCCTTAATGGGAGGGCTCCTGTGAGTGCATTAGGAAC  
TAGCCAAGGAGCCTTGCTTGCCAAAGCTGTCTGACTCAAAGGAGAGGAAGGGACANATGGCCTGC  
TGACTGGGGCTGAGGCANAACCTAGATTTTCTCTCTTGNNGTTTAAAGATATTTTAAATCTCGGAAT  
TCAAATCCTATAGTGTGAATATCTGGGGAGTTCTAACTTCTGGATGAAAAAGGAACCAATTTAGTG  
GTAAGAAATANAAGCCTGCTTAANAGGGACCTTAACCTGCCTCCTTGAGGAGTAANGGAGTCANAA  
GGAAGACCTAACTACCATTCCTTGGCCCAAACCANTTGGT'TTTACCCCATACTCCTTTCCCTG  
GGGGTTCAGGGACCCTGACACCCATTTANGGGAGTAAACCATTTANAAANCTCCCNCTTGCTGC  
ANCTAAAAATTTGCAAGTCAAATTCCTTTTTTGAATTTTAAAGCTTNANTTTGGGNTAAAACT

SEQ ID NO: 562 GGTACTTTTITTTTTTTTTTTTTTTTTTTTTTNGGAAATAAACATTTATNTA  
AAAATTAGTTTTGACNTTTTAAAGNGAATGCANACAAGNGTTTCCAATTNAAAAGNCCATTG  
NAAGCTANANAAGTAAATTCGAAGNTGGCAATAACTGACTCACATTNTTTACAAGTGGCCTANA  
CAATANGGAACCTTTCACCTNAAATTCACAGAGCCATGAATCACCTNTGCTTCCCCATGACCTTT  
CCATATCCTTCTACTCTGTCTTCCAACCATGACACAGAACTGAAACATACTTTAAAAATCTNATT  
CTTGCTAGGCACGGTGGTCACATNTGTAANCCCATCACTTTGGGANGCCAAGGCNNGCGGATCA  
AGAAGTCANGATATTTGAAACCACCCCGACCAACATGGGGAAACCCTGGTTNTACTAAAAAANAC  
AAAA

SEQ ID NO: 563 GGTACTTCTGGATATGGGATGAATTTTAAAAAGATCTGGGAAACAACAGTA  
AGGGGGAAAAAAGGCCAGAGTATGATTTTTTAAATGGCATCAAAAAAATAGCTATTTAAATATTT  
GTTATTACAATACTACATAGATCCAAGAAGTGAAAAATAACAACAACAAATCAGTTGACAATG  
CTGATTTAAGCATGCCATAGGCATCTTTCAATTATGACTATGATCACCAGGNCACCTGGACCAGAC  
AAATATCCCTGTTCTCTGGAACATATGCCTCAGGTTACAAAACTAAGGAAACTACCACAAAGAT  
TTTGAGCAGCTCTATTCTTGACCTGCACAATGCTCTACAATGAATATTATGTGATAAAGGCAAAA  
GTGTAAGTGAAGCTTTATTTATAGTTACAATCTTTGCTTGGATGAATAGTTGGTGGGACAACATT  
CTTTTAAACTATATAGTACCCATTGCAGGACATTCTCACTGGAGATCAGCTTCTTTTGAATACA  
AGCCAGAAGTCATTGGGTGAAGCCTTGAATCAGCTAGTTCCCTTAAAAAACAACAAATCTTGTTTTAC

TGNCNTNGGTGCTAATGAGGGGAAAAATGTGAACCTCAANGGANAAAAATGGGTTTTGGAAACTCAAT  
ATANNNATAGAAAAGATATNGGAAAACTCTTGGGGCNTGACCTGGGGGAATANGTAANTTTTCGNA  
ATTTAAATT

SEQ ID NO: 564 ACATGCTCTTCAGGTTCTAGGGCTCCTGTTAGGGGAGGGAGAAATGTTGAAT  
CAAGAGGGAAAAACAACACTACTATGATTTATAAACATATTTTAAATGTAAAAATTTGCATTTAAAAGG  
AGTGCCCTGTTTTCTGTGTTAAAACCCCATTTGGTGCTATTGAGTTTGTCTTTATCTTTTATCCC  
AGTGAAAATTGTTGATCTTGCTGTAGGGAAAAATTAACCTCTTTGAATCTCCAAACAAGGAAGTTT  
CAGCATTCCTTATGGATCAGAGGAACCTTAGAGACCTGAAATTGTTGCTTCCAGTTTAGCTGCCC  
CTCAAATTCAAGTGAATATTTTCCCTTCTCCCTTTACCCTTCTCCAGAAATAAAGCAGGTGACAGG  
GTTTTCAGAATCTTAA

SEQ ID NO: 565 AATTGCCCCTTTCGAGCGGCCGCCCGGGCAGGTAAGTCTAGGACTGGCCA  
AAAAATGGGCAAAATGTATCACTCCAAACACTACTGATTCAGCATTGTTTTCATGTCTTAAAAATTGC  
CACCTGCACCTTTGTTTCTGCACTATTATGTAGTGCATTTTAACTTAAATTTTTTCCAGCAACATGTT  
ACTTATTTAAGATACATTACTGATATTTTATTATAATTAGTTCACCTTCCCTGTGAAACAAGANAAT  
TGTAATGTTGNGGAAAAATGATACATATGTGGATGCTAATGAAATCATAGTATTTTGGGGTAGCT  
NCTCTGAANACCTNAAAGACCTGCGGCTTTGGTTTATAAGTGTGTTGGGGCCTNATCAAGCCCCATCT  
GATCCAANATCAATATTTTTTTGAANAAA

SEQ ID NO: 566 GGTACACAGGCGAGGGCATGGACGAGATGGAGTTCACCGAGGCTGAGAGCA  
ACATGAACGACCTCGTCTCTGAGTATCANCAGTACAGAATTTGGAGCTAAGGACTGTGACTGAAA  
TCATTTTCCCATATGAGCAGACCCTGTGTGTCAGGCTGTTTCCCATATGAGCAGAGCCTGTGTGC  
AAGTCTGTTTCTGGCATGTCCCTCATTGAGGAAGGGAAGCAAAAGCTGGTTATTGCCAGGCCTATT  
AACACTTAATATGCAAATTCTATCATNCTGAAACTGGGGCATCTGANGAAAAGGTGACCTTGCTG  
GATGGCTTTATTTGCATGGCTCTGCCTGTCTGCAGTGGTTGAGTCTCATNACCTGGTATTNTGTAT  
GAGCANATGTGTGCTGAACGTTGATGCCCAGGCAGATACCTCNTNAAGACTGCTGCAGGCANTGC  
TTAAAAATGANCCAGT

SEQ ID NO: 567 ACTCTAATTTCACTAACTGCCAAAAGGTTTTCCAGAATAATCTCAGTTGCTTC  
ATTCTTTAAAGATGAAGCCCGAAGAACGCATGGCGATTACTTTANAGGAACAATTAGCAGCAGA  
GGCAGGGCTGTGCTGATCCCATCTGGCATCGCTGGGAGCTAACATTAAAGACATGGCACTTTGGG  
TCCGGGTCCAGGTCCGTGGTTGAGAGCAGCTGCCACACCGTGGCTACTAGAGGATCCTTTTCCGGCT  
TTGGAACTGAGGCTGACTGCACCATCATCTAAAGGCCCTGAGACTGCTCGCCGTGCTCAACAC  
CGACTGGAGTGGCCATTGTCTTCCANCCACGCGGCCGACCTCGAAAAAGCCCCNCGTACCT

SEQ ID NO: 568 GGGTACCTTCAGAAGCTGGATACAGCATATGATGACCTTGGCAATTCTGGCC  
ATTTACCATCATTTACAACCAAGGCTTTGAGATTGTGNTGAATGACTACAAGTGGTTTGCCTTTTT  
TAAGGATGTCACTNATTTTATCAGCCATTTGTTTCATGCA

SEQ ID NO: 569 GGTACCAAAGGAGAATTTGGAGAGCTGGCTAAATTATTTGAAGAAAGAATTG  
CCAACAGTGGCGTTCANAGCCTCAACAAAACCAAAGGATAAAGGGAAGATAACCAAGCGTGTGA  
AGGCAAAGAANAATGCTGCTCCATTGAGAAGTGAAGTCTGCTTTGGGAAAGAGGGCCTTTGGAAA  
CTTCTNGGAGGTTTTAGGAACTTGCAGCAAAGCCATTGCGGTTGGAGTAATTGGTTTCCCAAAT  
GTGGGAAAAGCAGCATTATCAATAGCTNAAAACAAGAACAGATGTGTAATGTTGGTGTATNCAT  
GGGGCTTCAAGGAGCATGCAAGTTGTCCCCTTGGACAANACAGATCACNNTCATTAGATAGTCN  
ATCTTCATCGTATNTTCCACTTAATTNCTCCTNTGCGCTTTGCTCTTGCNAAATCCATCNCGTATT  
TGAANNATAAATTAANCCGTATGGGAGGNTGCCNAGTGNNATTCNTTTCCAGNGCTTGATTG  
CTCNGANCANGTAAGNTACCTGGCCCCNGNCCGGGTGCGNTCGAAAAGGGCCGAAATTACCAA  
CNCCANTTGGCGGGCCGTCTTCTATTTGGAATCCCNGCCTT

SEQ ID NO: 570 TCGAGCGGCCGCCCGGGCAGGACGCGGGGATGATGATGAAACAGAAAAATGG  
CCCCAAACCAAAAAACGACGTCCACCAAAGCAGAGAAGAAAAAGGCTCCCAAGCCAGAACGT  
CTGCCTCCATCAATGAAGGGAAAAATAAAATCCAAAGCCATAATTTTCATCAAGTGTGACTCTTC  
GGATGAGGATAAACTTAAAAATTGTTGATGAAGGACATCCCAGGAACAGCAACAGCAACAGTGACT  
CAGACGAGGACGAACAACGAAAGAAATGTGCCTCATCAGAGAGTGATTCCGATGAGAACCAGAA  
CAAGTCTGGCAGCNGAGGCCCGGCANTCCCCGGAGGCC

SEQ ID NO: 571 ACCTAGCTCTGAAAACACATCTACAGAAGCAAATGAACTCATCTGCAAAATA  
AAAAGCACATATCTTTAATTTCTAATGTTTTATTATAGATTTTAAAGATACATATTTATTTTATATT  
ATTAGCTTAAAGAAAGTAAGTCACACAAGAATAAGCTTTTGATGTCTGCTAAGGTAAATTATTAT  
ATTATAAGTTTGTATGAAGTTATAAATAAAACACATATTGTGCTTATGTGAATAAAACCCAGCNT

AATCGGAGTGCATTTGAAATGCCAGAGTATTCTCATTGTTTCANTAAAGACAAAAGAAGAACATTTTC  
CTTTTAGAATTAAANCTTGGCTTCNGTTTTCTTTATCAAAAANAATNANTNTGAATACCNNTTTG  
ATTCAAAACNAAAATTTGTAAAAANNNAAGGNGACTAAATNAAACATGTTCTGCGCCTTNATT  
NAANNANCCATTTTNGGAANAACCTGGAGGCAAANACCAAAGTTA

SEQ ID NO: 572 ACTAAGTGTATACGTATTTTTGCCACTTTTTCCTCAGATGATTAAAGTAAGTC  
AACAGCTTATTTTAGGAACTGTAAAAGTAATAGGGAAAGAGATTTCACTATTTGCTTCATCANTG  
GTAGGGGGGCGGTGACTGCAACTGTGTTANCAGAAATTCACAGAGAATGGGGATTAAAGGGTAGC  
NNANAAACCTGGAAAAGTTCTGTGTTAAGANCTTGCTGGCAGAAATAAACTTTTTGGAAAAGTTT  
ATTANNCACA

SEQ ID NO: 573 ACGGTCTCACAGACAACGTTGAGAGAATAGTAGAAAATGAGAAGATTAATG  
CAGAAAAGTCATCAAAGCAGAAGGTAGATCTCCAGTCTTTGCCAACTCGTGCCTACCTGGATCAG  
ACAGTTGTGCCTATCTTATTACAGGGACTTGCTGTGCTTGCAAAGGAAAGACCACCAAATCCCAT  
GAATTTCTAGCATCTTATCTTTTAAAAACAAGGCACAGTTTGAAGATCGAACTGACTTAATGGG  
AAGAACAGAAAAATTTAGTTGCTACTGTAGATTTTCATGATTNAGANGCAGCCTTAATTTGCCATT  
GATCATNCCNTNTTTTTNGNTNTTAANAANCCTTNCCGGCCA

SEQ ID NO: 574 ACGCGGGTGTGCTTCCTTTCAAAGGGTTGGACCTTTAAATTGCTGCAAAAGGT  
AAATGTATTTTTTTTTAAGTATTGGTGTCTTTACTCTAGCTAGGCTAAAATTTGCTAAATGCCTT  
GGTTTCTTTAAAAGTTTCATGTAATATTTCTGATTTTTCAGAATATTTGCAATAAGAGTCTGGATTTT  
AAAAAACACATGCATACACACAATTAAGAGCTCATGTCTTAGCAAGATCTGGGAAACCAACATTG  
CGAGAGTAGCTATTTGAAAGAATAATTCTCCAGAAGTTAACATCTAATATCTAGTATCACCAAAC  
AGTATCGCTGGTCTTTTTATTCAATTGGAATGNATTTATTTTTTACCTTACCATTTGTCCNATTAA  
ATTG

SEQ ID NO: 575 ACGCGGGATTTCGTAACCTATTGTAATTTCCACTTTTGTAAATAATTTTGATTTA  
AAATATAAATTTATTTATTTATTTTAAATAGTCAAAAATCTTTGCTGTTGTAGTCTGCAACCTCT  
AAAATGATTGTGTTGCTTTTAGGATTGATCAGAAGAAACACTCCAAAAATTGAGATGAAATGTTG  
GTGCAGCCAGTTATAAGTAATATAGTTAACAAGCAAAAAAAGTGCTGCCACCTTTTATGATGATTT  
TCTAAATGGAGAAACATTTGGCTGCATCCACATAGACCTTTATGTTTTGTTTTTCAAGTTGAAAACCTG  
CCTCCTTTGGCAACATTTCTGTAATGAANCANAATTTTTTTNTNTTTTTTNTAANTTTGTAANTTTT  
GGTTCTTGGAANANGGNTCAATGGGTATTNGGGGCTGGGGTATTGAACACGAAATTTTATTAN  
CATNGNGNTCAAAATATACCANTGGTNAGGTTTTAAAAAGTATTCTTGANGGGTCTTTTCTATTT  
ANTAATTNAACTTTCAATAAANNGTCCCTCGCCCGGACCACN

SEQ ID NO: 576 ACGCGGGGCTTTTTTTCATTCCCGTTGTTATGGAGGGCCACATCTGCCAAAGC  
CTGGAGTCTGCGAAGGCCGGGACCCGGTTCCCGGCCCCACAGTGGGGGTGTGCAACCCGANAGA  
ACTGGGTGCAAATTCGTGAAGAATCAGCATCATGTTTGGCAGCTGAGTATTGGAGCCAGNAGCC  
TGCCATGANGTTTTGAGAACANAGTGCTGTTTANANCTGGCAGCAGCATCTCAGCCCAANAGAA  
GGTTATATTCCCAGAGGATGTCAGTCCCAATGACCANNAGCTGNCATTAGATTTGGATTCTGAAAN  
TAANNGGCTTAACAATTGGGTGNATAAAACNTGNNTTTCTNTNTTNAAGGANTTG

SEQ ID NO: 577 ACTGCCTAGCATGGTATCTTCCTCATCATCAGAGGTGCTCTATACATCTTCAG  
TTGAAACGTGAAACAGACATCCATATGCAAAATATTTCAAGGGCTTTGTTGGCTTTTACATGTTT  
TTCTTTAGATAACTGGTAATGATGCACATTACAAAGGAGACTTTTCTAAATCTCAAGTCCTTTGCT  
AATTTTTCTTTGGAACAACGGCACATTTTCAATGCCAAACCTTCTCCTACAACATACAAAGGGGAG  
ATGCCAAAACCTCTGAATTCCTGTAAACGGATCCTGCAACTAGTTCTATCCAGAAGATGGAGACAATA  
TTCCCTGGAGTTGACTGAACATGTGAGAAGGCACAGCTCANAAGGAGAGGAAGGCTGAGGGCAG  
TGAAATGAGAACCTATGCATCACCTGGCCTTTTTACATGTTAGTCTATCCTACTATCCAGGAATTC  
ACTTCTGCTGTACC

SEQ ID NO: 578 ACGCGGGGGCGCTACTGCCGGAGCGGGGCGGTTATGGCGGGCGGAGACTGCG  
GGCCCGTAGCTGGGCTCTGCGAGGTGCAAGAAAGCCTTTGAGGTGAAGGTGATGAAAGTCATCA  
TAACAGATGTTTTCCAAAACTTGTAGAAGGTTGTAAAAAACTACTAGGATCACGCGGCATGTA  
TTGAGCATATAGGTTGCTGTAGATGAATGTTCTTAGCTGTCTGTTTAAAAATACTTCTGCTTCGTT  
ACCTCAAGTGTGGCATGCAGCATTTTGGAAAGGAAATGAAGACGTGTTCAAGAAAACATGAACA  
GAAGCAAATGATGAAAATGAGCATTTTACTTGATGTTGATAACATCACAATAAATTATGCCNANA  
AAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 579 ACTGTANAATGTGATGGAAAAGCATTGATGAGAATTTATTGGCAGNTCAGAT  
TGTTGTTTTCCCAACTTAGTCTCTTTATTAATTGGNTAAGGTTTTCTCCAAAAGGGCATTTCACAA

TGGGAATTATTTAATGTAACAGTGGGCACAGATTACTTATCTTCCTTCTCTGCTTTGTGACTCACCA  
GCAGTAACACACACAATCCACATCTTGTGCACCTCAAATGAACAGACTTGGTTTCCTTGCTTTCTT  
GACATTTCCATGACTGTTTCACATACAACTATTGGGTGAGGTTTTTCAGCTGTTACCGACCCACG  
TCCTGCTGTCTGTGTGGCCTACAAAACCTGTCCATTCCCACCCCTTGTCTTGGCATTGCAAGA  
GTCTGGAATTGTCAGGTCTCAGCTTCNAAAAGTCTGGTTCCACTGACAGGACACATTCTTTAGTG  
GGAATTAGACCTCAAAGTCTAGTTTTGTATGTNGGTATGAAGGGGAATTTTTTAAATAAANTTGAA  
AAAGCTGTGAACAGCATTAGAACT

SEQ ID NO: 580 ACGCGGGGAGTTGTAGCTCAGCGTGGCTACAAGTAACTGTGGTGTGGAAGCA  
GAGTAGAGAGAAAACCTTGTTCCTCATTAGAGAGAGAGCCACACTTCTCACTGCTCACAATGAGAG  
GCCAAAGATTACCTTGGACATCCAGATTTTCTATTGTGCCAGACCTGACGAAGAGCCTTTTGTGA  
AGATCATCACTGTTGAAGAGGCAAAGCGCAGGAAGAGCACATGCAGCTACTATGAAGACGAGGA  
CGAAGAGGTGCTGCCTGTCTACGGCCCCACAGCGCTCCTGGAGAATATGCACATCGAGCAGC  
TGGCCCCGACGCCTTCTGCAAGGGTGCAAGGGTATCCATGGAGACTGGCCTATGACACGTTAGAG  
CACGGGACCAGCTTAAAGACGCTCTACCGGAAATCGGCATCACTAGACAGTCTGTCTATTGGT  
CATCAAAGATATGGATAATCAGATTTTGGAGCATATGCAACTCATCTTCAAGTTCAGTGACCA  
CTATTATGGCACAGGCGAAACTTTTCTCTACACATTACGCCCTCATTTTAAGGT

SEQ ID NO: 581 ACTTCATTTTTCTGTGGCACAAGATACTCTAGGCTCATCTTGTATAGTTCATAT  
CCCAGCTTAGAATCAGTTCATTTTCTAAGGAGCCCTGGTTCCTTTTATTGGAAACCAAAATCTGG  
GCACCGGTGTGTCTCCATTCTAGTTGTTTTCTGACCACATAACTGCTAACAAAGATGCTTCACTCT  
GGCTACACTGATGTGAACCTTGAACCTTTAGCAGAAGAGGCTCAGCTTAGAGAACAATGAGCTCCT  
ACATTACCTTTTTTCTCAAAGAATAAGTAAGTCTAAGCAGAAAAAAATATGCAAAGAATTTTC  
AGTATGAATGAAATAAGACAAACCATCAGGCTTGTGTATTGTAAACCAACACAATATAGTTATA  
ACAGATCTGTAGAAGGATCCTTAGAATAAGAGAGTCATTTGTGGGGGTCTCAGGGAGAATACT  
GATAGTATCTTCGGCTTTGNCCGCATAACAGACACANCATGGGATACTCCTGAAATTCATCCATGA  
TACTGAATGNATACTCTGTCTGGGCTATTACATGAAGGANCAATTCTTA

SEQ ID NO: 582 ACCACAAATGCAGAATCAGAGCAGTAGGAAGAAGGTTAGTGCAGTTATACTT  
TCATTAATAAAAAATTCTGAATCACTGCTATTTAAAAACACCTTGAAGCAAGTCTTTTGTGTTGAGAT  
TGTTTTTTAAACTAAGGTAGCAAACATTTTGCCATGTAATGGCAGTGTTATATGCCGTTATCTTGCT  
TTGTATAAAGAAAACAACATGAGAGATTTTAACTGAGGTTTGGTTACATTACATATTTAAGCT  
TCTACACAGAATGATGGACACTTCGAGAAGCTAATCCTTATCCAGAAACATTTTAACTCTCTAAAA  
AACAAAGCAAAACAAACAAACAAAAACCCAAAACCTACGTTGCTCCTTTTACAATAGTGCA  
CATTTTACCATAATTTAGTTATGGCTACAAAACATCAGAAGATTTTTTTAATGTATCTTCTCTAT  
GGTAATTAATAAAAAAAAAAAAAAAAAAAGT

SEQ ID NO: 583 ACGTGTGGACACCCTTTTAAAAAGCAATCACTGTGCTAGAAAAGTATATT  
GGCTTTGTAGGATTAAAGTTCATTAACCTTCAATGTAATCATGCCTCCTATTACTGAAGTCAGATTG  
GAACCACTAAAGATCCAACTTTCTGTCTGGTAATAGAAAGTAAAAATCTAGACATCATTTACATT  
TGAGAAAGCTGTTTTTAACATTATTTTAAATGCCAAATATGTTCTTTCTAGAAAAATATTTATTTT  
TGTTTTTGTGGATAGCTTTTAATTACATTTAGAGAGGTGTAATTTTGGGTAGATGCTCATTACAT  
TTTTGAAAGGTTATGATTCAAAAATAAGATTTATATGACTGGTGATACTGGCTTTACAGAAATT  
TCAGAGAACTAATTTTAAATCTTTAGCATTTAAACCTTTTTTGTGTTGNTTCTGACATATTCTG  
ACAAAAGAGCAGCAAACCACTGCTGTGTGGCATTCTTGAGGTGTGCTGTGAATGTGCTTTTAAAG  
AAATTAATAAAGAGATCCTTCTT

SEQ ID NO: 584 ACTGTAGGTATTTATTAATAATAGCAATGAAGATGAAAGAGTGATGTATCAG  
AGAGGTGGAGATAAAATCAGTAAACCTTAGACACTAAATGATAGGGGAAGGTGGAGGAGAGGAA  
TGAGCCTAGAAAACCTTAGAATATAATGGTTCTAAAATTAACCAAAGTAAGGGACACAGGCATTAG  
AGTAGGTTTTGCAGAGAATGAATGTTTAAAGACACACACAGGTGTCTCTGGGACAACCAAGAAAA  
GTGCAACAGGCAGATGGATTGAGGAGTCTGGCTAAAGATAAGGATTTAGGAACTGCTGAATTA  
ATTACCCAAGCGTGAGAAAGTGGTGTGATTAAGAGAGAAAAAAATGGAGGTCTGAGGAAT  
AACATTTAAGGAATAAATGAAGAGGCCAAAAGGTGGGGGTGGTTCANGAGTGANCAAAATGTA  
AGAAGTCAAGGGAATAAATCTTTAAAGNAGGGGTGGCAAAAAATGTCNAATCCAAAAAAAGN  
CAAGTCCCTCGGCCGNGANCCCN

SEQ ID NO: 585 ACTTNTGAAATGCCATCANCAGGCCTCCTACAGGAGTGCCCANCAAGGCTCC  
AATNATGCCACCAGCCACCAGGCCACNAGGCCTACGNTTATCCTAAAAAGACTTCCCGTGACAG  
CTCCTGCNATTACAAAATGGCTTAAGGCATCTTTATTTCTGGTATACATNCANACTAGTGTTNACTG  
TGTTGAATATAGNNACAAACACTGCNNTTCTCCAACCCACGCGCCNGCCATAACNAATGAANNCT  
CATGTGGCANCACTATGTGCATATTGCACAGCATCNAACCGGCTTGATAAATTTCTGCCTGGCTC

TGCTNAATGTATTGTTGTNTAGCATGAATAAAAAGCTGGTATTTCCCATACACCNAATCAATGATGC  
CTGCTGTAGCTGCNNGGNTACACAAAATTAGCAAGGGGNCCTNTGAANTTCTTTGCTGNTATCNAT  
GCCAANCANCTCCCTGNGGCGNGNTCCNTCCATATTCCGGATAATAGGGCTTTGGNACGTA

SEQ ID NO: 586 ACCAAATAGGGNTNCCCACCCACCCCTGCGACAAGTGCTCTTCTAGAACAG  
GTTCCCTACCAGCAGCACTGGNGTGAATGAAAGAGAGACCCANCCNCGTCTNACACNNGTGGAATT  
GCATTCTTANCNAAAANGAACTTTATAAAAANTTNGGGATTTTCTTCTAAACATAAAAAATANCC  
CC

SEQ ID NO: 587 ACATTTGCCAAGACAAAGGTTTCTAGATATTATGATATTC  
TGAAATAACTATTTTATCCTGTAGTTCTATGATTATATGATTTGTAAATAAGAAGCCTAACCAATT  
TAAAATTCTGACTTTAGTCTCATATATCTTGTGCACTGTCTCGGTTCAAATTCAAACCTACCAT  
CTTCAGTTGTGCGACCTTGGGCATGCTACCTAATCTCTTTGTGCCTCAATTTCTCCTTTTAAAAATG  
GGGATGATGATGATNACAATAATACCTACCTCACAAGGGTGTGTGAGCATCAAATGAGATAACA  
CAAGTAAATGCATAGAACAGTTCCCAAGCACAGAGTAATTCAAATAAAATATTAAGTAGTNATAG  
TAGNGGNGNTAACTCGNGAATCTTTTAAATAACATAATANGCTTTGATTTTATTNTCTCTTTAAGTT  
GTTAACTTCTTTTCCCTTGATAAGTTTTATNTCAAGTAAGGTANGGTGTTAAGTTAGTTACCCAT  
GTCCCAATCAAGGGACCTAATGAAGATATATATTAATAATTNCCTTTATTGTTGG

SEQ ID NO: 588 GCCCACTGATGTCCACATCCCANANAAATTNTCAAAGACGGAAATCAGGGC  
AAAGACTCCATATAACAAAGCGCAAGGATATGCTACTAAAGTTGCTGTCTTCCATGGGCTAANG  
CAGAAATAAAATATTTTGAAGCAGAAAACTACACCATCCAATAATCCNCGAGTGAGAATGATT  
CCTACCATTCTTGTCAAAGAAAATATCACTGCANAGCTGGAAAGTAAGATCATGGGCAGAAGACA  
ATATCCAAGGACACTTGCCACACNACCNAATGAAACACCTGNCATACTCATTAAGTTTNATAAAC  
AAAACATTCTAGACATTCAATTGCACTGATCCCCGNNACATNNNCAAACCTTGGNTTTTGCANNC  
CAGTATCATTGTGGCTCCNNATGCANGCAAAANNCCNTTGGACCTGCCAAATTAGTT

SEQ ID NO: 589 ACTTGAAGATATCTCGGCAAACTGAAACAATGGTGACTTCTGACGTAATTCA  
GGAGATTCTGAAACATCAGGCCTGCATGTGCCAGAGGATACCTCTGGTGTGCCAAGAGCATGTTT  
TCCAGCTAACAGCAGCATACTCAGGCCTCCCATTGAGTAGGATCAGCCATTCCAAAGTTAAGCTG  
AGGCATTTCTTCAGTCTTTGCTTTCTTGGGCTTGGCAAGTCTCTTGAAGAGTCAATGGGAGGAGGC  
CCAAAGTTTCTTTCGTGGATTGACAGTGGGTGTTGGGGATTTCACANGCTTGTGTTGTGGTANGACA  
CCATTTGTAGCCAGGATTTGCTTTTATAAATGCATNCTTANACTNCTTGCCATGTCTGNGTNTTNT  
GNTTTTCTTTGGGATCAAGGAACNGCCNACCCAATCNGCTAGNTCTTGGTTGCCCTCNGGTTTC  
AAGCTTGGGGGGGTTTCTTACGTACCTGCCNG

SEQ ID NO: 590 ACATCTTAAACACGACATTGACACTGCAAGTGCTTGGATGTCTGGCACCTCA  
GAACACAACCTTTGTGCAAAAGGAGTGCTGTGCATCTGAATAGTGCAATTCATTTCCAATGCAAGA  
ATGCCAGTTTCTTAGGGCCCCCGTCAAGGATATTGCCACTTGGTTAAGTTGTGATTTCAAATCCC  
ATTTTCTATGGGAGATACACTGGTGGAACAAATCTGCAGAGGGCCCTGCACCTGCGCTGGTGGAT  
CAACGGTGTCTTCTCACCTCCAGGGCTAGCTGCTTGCCTTCCCTCGTGGGGCTTAAGTGGCTGGGA  
AGCAGATAGACAGTGTAGATGAAGTCTGGGCCCTGCTGCTGGACTGAGCAGTGAGATGACTGGG  
CTGGCATTGTTCTTTCTGCTGTGTGTGACCTCTACTATGTGCCTGGAACTGAATTACATGTAA  
TTAATGCTGCAACAGGATTTTCTTATTGCCTAAGCATGAACTTGGNATGAAAATAATAGTTAATA  
ATTTACTATGTAACCTGGATCTTATAACCTGGACTGTAAGTTTGTGTTACCTCCAGTGGGAGTTTNT  
TTAAAG

SEQ ID NO: 591 ACAAAGACGCAAATTTTCATAGTGCCTAGAAATAGCACAGATCTATTCTACT  
CAAGATTATTTGAATTTTTCAGGGTATTCTACCTAGAGCCTGTGGTTAATCGCCTCCCTGCTCCCC  
CTACCTTTTATCCCTACCCCCCTCAGGGAATTTGGATACATGTGAGGAATAGTCTTTGTTTTCTT  
ATGAACCTAGAAAATTACAGATCATAAAATCTGGATATTAAAGTAGTTTCCAAAAGCATCTCATG  
GGAAATCAAAGTGCTCGGCATTTCCGAGCTGGAGAATAAAATCAAGAATCCTTAANAGAGAAAA  
GAAAATGGAGATGGAAATATGAATTTTAACTGGGAAAAATAATTGGAAAGCTAACCTTTTCAAAT  
ACTCTAATTATGAAAGTTGGAAAAGGTTGTGATGCTCAGGATTATCCTGTAATGTGTGAAGATACA  
TAAATTAGCACCTAATTTATAGGCAAATTTTGGTAAAACACAAAAAATTATGGTATGTCTAANTTC  
ATGAAAAGTATGTAAATGCAGTGGCCTAAAACCTGTCAACTGNAATACAGTTAATGTGGAACCT  
TTCTAATAATATTGAGCTCTTG

SEQ ID NO: 592 ACACACCAGTCTTTCTTTGGTGTTTAACTAATGGGCTGAGGAAGGGAGGCA  
GCCAGGAAATTATACCCATTAAAGTGTGTGGTTAATGCCTATATATATCATTTTACTCACAGTTA  
GTGGACTTATCTTTTAAACAGAGGTATGTGGCAGTGGGGTAAGAGAACTTCCCTTATTATAGAGA  
AGATTCCAAGAATAGGTTTACAGATGATTTGAAGAACCATGTAACCTCTGCTGCCTTGCACGGTCCC

CTTTTtagcactaataactcctaagtgctgataatttcaagggttagcagcactcccaaagaagct  
gtccgggctgggtgggtgggtggtcctatgcctgtaatcccacaatttcagaggctgangcgggcgga  
tcacctgangtannganttcnnagaaccagnctngcnnccangnggaaaccccntttcnctaaaa  
tacaaaantnanccagcgtgggggnagggccctgtattcccagctacttgagagctggaggnng  
ataatngccttganccagnaggcagagtttgntnngagcnaaatgcgccattgtttc

SEQ ID NO: 593 ACCCCAacttttgctggacctcatgcagcttttagctaataaaaagtttctttaag  
gcagataaaagttagcaatgctgtggaataaaaaagctactgctgtgttggttaatagctagcacaga  
tgttgacaagacaggagcttcctactatggagaacaaactctacactacattgcancaaatggan  
aaagtgcctgtagtgcaattaccaaaaaatggccccatttatgatgnaagttnggaattntagttct  
actgagatntgtgctgtatattggntttatgcctgcccgaanngacnatttttnaactttnnantgngat  
cctgtatttgacttttgganctgggncnncnnaatgcaagncctantatancccttgganataatttt  
taantnttaacntggtnntgnaaanncaggggnnaaatgnaatggtngggntgggtggaaaaac  
tncaannttttntaataccggnntggttatgnattntacctnatttgctgtgtaacc

SEQ ID NO: 594 ACATACccaaaagaattaaaaagcaaggacttgaacaaacatggctcctagca  
gcataatcacagtagccaagagggtggaagcaacctaaatgtccatcgacagatgaataaaatcaa  
caaaatgtggaatagtcatacaatgaatatcagccttaaaaggatggaaaattctgacacatttta  
caacatcgataaaaacttgaggagcttagtaagtgaacagggccagatacaaaaagacaaaata  
gtgatagttccccctcagatgaggcacctanaatagtcaaatccacagacagaaaagagaaatgg  
aggtttcagggggcaggagagagaattgagcggttagtggtttnntgggtgcagattttcantg  
ggaagatgaanagggttccaggggttncccgacnacaagtgggnatnntactttnaaaatggctnn  
aganggcaattntatnttagggattttttnaaacacaganttaaaaaanaaaaagntttngtttag  
ggttngcangancnataccaaanancnggttttccnnttcngcgccctg

SEQ ID NO: 595 ACTCCagcagcaggagagcggatttacaacatctcaggggaatggcagccctc  
ttgctgacagcaaaagagatcttccctcactgtgccagtgggcgccggagagagcctgcgattattg  
gccagtgaacttgcagaggcacagcattgcccagctggatccagaggccttgggaaaacattaaagaa  
gctctccaaaccgtctcgcccaaatctgcagcagcatacggaccacaaatgagacaccaaaagttg  
acaggatggacttttaattgggcaacttctgggaccctgaagagacttcttcccttcaggccttattgt  
ttgagtgtagaagttccagagcaaggagccatgttcccttaagggaattcaggaattcagacgtgc  
tattcccacaccagttaggtagagctgtctgttcaccctcccacagctgatcccagctcactgc  
ttgctggggccatgccatggaagcttccatcagctcagctgaatcctcctgctctctganctgc  
tggcttttgccctnctgcaactaacatnctcttaaccttgccctgccttgcat

SEQ ID NO: 596 ACTGACacattacgactcttggttggtgacaaacatttaagcagggctatccaa  
agctcacagtcttgaaggcagggaggtgtgtgtaaatgttaggtttttaatttaagcagtaagcaagc  
tgagatttttcttggtaaaatagtttgtttattcatgtagatctaaaaggctggctcagtttgtcag  
agcacaggagaatgccagaagaggtcaaggccatggggatgtatcgacatgaatccatcccaac  
ttgtggaacaaacgacacagcagttaccatcacaaagaactctattagtaggaatagaagact  
tgagggtccaattcagttctcatttggattctatagtatctctaagaatttggttaaaaaaacaaaa  
caaacaaaaacaaaaaagggaatccttactcctttcattaacaactgccacaaggagaagaaat  
aaaagacaactaatatagatnngccttctgattaaaaaatctgtgagcttgaatacatttttaatat  
gttggaatctagagtgggnaaccaatgggttaataatactacaaaactttta

SEQ ID NO: 597 ACTTTTctttttcttcccttttttttggaattattttcctgagcctttttgtttta  
cggatatttgtaaaacttttatgttaaagaaaaaatatacatttacaaattgtgagatttttaagag  
aaattttctacgatgtatactggcttattttttaattttaaaacgggggtttccgtcgccactgggtgga  
gggggtgcgccgttagtccccctcgctcctggctttgggggttgaggacttgnggtccagaaactct  
gggagcttctagaagaaatctactgagngtatttctgttntttgtnaattccttgctttgtcgaactg  
acctgcnnngnnggtctgangtgaactgnggggntgngcacagnacagccgagnggatcccncc  
cagcgtganccntnccnagatggaangcctttctnccaaacccnngcctnngngggcgngtccg  
ccattnaccactcttgccacttggttn

SEQ ID NO: 598 TTCGCCcttttcgagcggccgcccgggcaggacgcgggaacgacattttttgt  
aaactttacacttttttggttatttttatttttaaaaaaatgaaaaatttaatttaaaaaaatgcaaaaa  
actgttggaattattttatnanaattccccctttgtgttggaactgcaaatgagtttcttctctt  
taggcctttcacaaactaggactgagaatgtatgtaaaagtctgtgacagctacctcgcccgcgacc  
acgctaagggCG

SEQ ID NO: 599 ACCAAGtgagtggaatacatattctagttaaagcattttgtgtctagctacac  
accgctaacaaagttacttagttatcaatgtaggattcttaaggagctttaagctaaggaaacctt



TTAGTGACTTAGCTTATTTTGTATCTTTTCACTTAGGAAGATTTTGGAGGTGATTTTTTCCATAGG  
 AGGATACCATCTGGCGGCTGCACATTGTAACAGTAAAGGCAGAAAGCTGTAGTGATAACCTCTCT  
 CCTAAAAGAGTTAACTGGTCTCATCCAGCAGAAAGCTATCTTAAATCTGTGATGTGTGTCAGGTGCAGC  
 CAAATATCACACCTTCTGATCTTAGCCATCCCAAACAGTATCTGTCCCGAGAGGAAATTCCCCC  
 ACCCCAGAAAGTTTACAGAAAACTGCCCTTCAAGTGTTTGCCTTATTCAGCTTTTACTTGTGCCA  
 TTAAAGCAAGCACTGTAGCAAAAGCCCACTTCCACATGGCCCTGGCAGGGAGCACTGCTGCTCCA  
 TTGCTCATTCTCNCTGTACCTCGGCCCGCGACCACCGCTAAGGGGCG

SEQ ID NO: 600 ACAATATGCAATAATAGTGACATCTTTCAGAAAGAGATGTTGAGAATGGGAA  
 TTCATATTCCTGAAAAAGATGCTTCCTGGGAATTAGAGGAAAACGCTTATCAAGAGCTTCTGCAGC  
 ACTATGAGCGTTGTGATGTTTGAAGATGTCGTTGCAAAGAAGGGCGAGACTATAATGCACCTGAT  
 AGCAAATGGGAAATAAAGCGCTGTGAGTGTGTGGTTCCAGTGGCACACATTTAGCCTGCTCCTCA  
 TTACGGTCATGGGAGCAAAATTGGGAGTGTTTGGGAATGTAGGGGTATTATCTACAATTCAGGAGA  
 GTTCCAAAAAGCCAAAAACATGTATTACCCAATTCTAATAATGTGGGGATTTCAGATTGTTTTGTTG  
 GAAGAGTCATCNCCTANATTACCCAGACAGNCNCCTGGATCCAANAGTAAAATCTNCTGNNGCNA  
 GGCAGCAAAATTTAGAAAAAAATGTNTCNACANCTNTTTATATTACNTNCGG

SEQ ID NO: 601 ACCAGTGCGAATCATCGGGCTATCCAGGTCCGAGATCCTAGTCTCCTGTCCG  
 CTCTGAGGAGGATGGATCCTTCTGCGGATACATGGGACCTCTTCTCACCTTTAATATCATTATGGA  
 TAAACAGGTTTTACATTTATTTGGGCTTTGCTGTAGCATTAGCCTTTGGATTTGTGTCCAGATTGT  
 CATCAAGAGCGCAGGGCAAGAACTTACAGGAAAAATCTGTTCCAAAAGCAGCTCAGGATTTGATGA  
 CAAATGGTTATGTCTCCCTTCAAGAGAAAAGACATCTTTGTGTCTGGAGTGAAGATTTTTATGGTT  
 CTCAGACTGGAACAGCGAAGGGATTGCGAACAGTTCTTGCTGAAGCAGTTACATCCCTGGATCTG  
 CCTGTGGCCATTATTAATCTAAAAGAATATGATCCAGATGATCATCTGATAGAAGAGGTGACTAGT  
 AAAAATGTCTGTGTCTTCTGGGTGCGACATACACTGACGGCCTACCAACTTGAAAGT

SEQ ID NO: 602 ACCAAAACCTGCAACAGGCTCATGGAACAGAGCCTAGGGATCTAGGAGCAT  
 AGGAGGTGGTGGTGCTGGGCAGGGCTCTGCATCCCCCTTCTCANCACAGCACCATCTTCACCCCTC  
 CTGGGAAAGCAGCATTGGAGCCTACACCGCTTGTGCTTTTCTCACCAGGGTAANAAATGCAANGTA  
 TTTGCGAGAGGGAGTGAGTCTGGAAGGTGGCAGCAGCAGCTAGGGCNAGACTTANGGGAACTT  
 GTGGGAAGAGTAACTGTGGAACCTACCTATGCTCTCTTGACCCCAAACCTCCCCAAAACCCCTCACN  
 TGAGGACTGTCTACCCCGGGGCTCAAAATAAACTGCTTACTGGAAGATGGGTGACTTAAAGGCA  
 AAANGGAANGCTGNCCCCTGGGCTCCCCAATCCCCTGCTTGCAANANCTGGTTTGTGATNCTNNG  
 AAAACCCCTGCATTTTNCCTTTTACGCCAANCTCCTCANAGNTTNANACANAAAGGGGNTGGAGN  
 NGANGGTCCTATNTNACTTTACCCCTCAANGCCNNTTNCACCCCTTNAC

SEQ ID NO: 603 ACTATATAAAAAAGAAAAATATTACAAGACCTTTTGAGGATCAGACATCACTG  
 GAATTCTTTTCAAAGAAGTCAGATTGTTCTTTATTCATGTTTGGCTCCCATATAAGAAGCGGCCA  
 AATAATCTAGTAATAGGTGCGNATGTATGACTACCATGTGCTGGATATGATTGAATTAGGTATTGAG  
 AATTNTGTCTCTCTAAAAGACATTAAGAACAGTNAATGTCCTGAGGGAACAAAACCCNTGCTGAT  
 ATTTGCTGGCGATGATTTTCATGTAACAGAATNATNATNAGAAGACTAAAANGNCTNNTTATTGATT  
 TTCTTCANAGGCCNCACNGTATAAAATATCCNNCTGCCNGGNTTAGAGTATGTTCTGCANTGAAC  
 GCNTTGNATGGGNGATTCACTTCGATCT

SEQ ID NO: 604 ACTGGAACAGGGATAAGTTCTTGGATAAGGTGCCAACATACCTATAAAAGCT  
 GATTTTTGAGTAAATTATCGATTCTAACATATGTAATGGATTTGGTGTGATAATTTTCTGATCTTTA  
 ACTATAAGTGACTTTTTATTCTCCACCAGAAAAAGATAAATGACTGAGAATGTAAGTCTGCGCTCTG  
 ATTAACACAATGGAGAAACGGAAAAACTATCTCTGTTAAAAAACTGATTCTCTGCTATTCTTCTGATA  
 TCAAATAAGAGGAAGGAAAAATAAACTTTTTGTGTGTAGATAGAAAAACATACCTGAGGCCAGGTG  
 CAGTGGATCACGCCTGTAATCCAGCACTTTGGGAGGCCAAGGCGGGCAGATCAGCTGANGTCAGG  
 AGTTCGAGACCANCCTGCCAACNATGGTGAAATCACGTCTTTNCTAAAAAATTCAAAAATTATCTG

SEQ ID NO: 605 ACGCGGGGCTTTCACATTTCGGGAAGCGTCGGGATTAGGTGAAAGAAGCTGAG  
 CTGAACACATTACGATGGATGATGGAAACATAAGACTATCAAGAAATCCAAGTGGTAATGGGCGA  
 AGTTTATTCAGCATCCGGCAATGGACTTATCGTAGTTGGGGAAACGGGTGTTCCGAATAATATCCT  
 GGAAGTTATCAGGACACCTATTTTAAATATAGGCCTGAATTTTGTNAAGTAATATTTAAGGTGGTC  
 CGTGATAATTAATAAAATGCTTAATTCATGTGACTAANAAAAA

SEQ ID NO: 606 ACGGAAGGATGCTGCAAGCTGACCCCAATAAAGTTTCTGCAAGGGCGAAGA  
 AAAGAGGCCCTTCCTCAGTTGGGGACCCCTGGGAGCAGGCAACCATTATGCAGAAATNCAAGTTGTG  
 GATGAGATTTTCAATGAGTATGCTGCTAAAAAAATGGGCATCNANCATAANGGACAGGNGTGTGT



GATGATCTCANGNGTAAGCAGATGCTTNGTCCA

SEQ ID NO: 607 ACATTCTTCAAGCACAGGGGGCCATCAATAGTAAAAATCTGCCCTTCACCATT  
GTTTAAATCTATGAGACGTTCAAGGCATTCCAAGGAATCGGTGCTCTGGAGCTCCTCCAAATTCGC  
CATGCACATGATATATAACTTAAATGGGAAAGGAAAGGTAGTGGAACCTGTTAGCTCTCACTT  
CGTTGATTCTNAGTGGNTANGGAATGTCCAGNGATNNTTTATNATTTGCCGANACATCTCTAT

SEQ ID NO: 608 ACATGAGAGTCCTGCCTTAATTTTGCTGTTTGGCCTCGGATCTCTGGGTTTGAT  
TTTTGCGTTGATTTTAAACAGACATAAGTATCCCTTAACCTGTACATGGTAAAGCATAAAATCGCA  
CTTGAGATTCTGAATTATTTGGCTCCTCCATTTCTGGAAATTGAGACTCAAGCTTTATGAATTTATC  
AAGAACTTAAAAATGAAGAAGGTCACAGATTGATCTTTTATAAGACCTTATTTTGATGCTTTGTGC  
TTCAAGGAGATGATACCTGTCTCATCATATAAGCAAACCTTTTGGCTTACAACCTATTTTTTAAATATT  
AGCCTTCTAGTCTGTAATGGAAATTGTATATTTTGATAGAAGTTTCTCCATTGGTTAAATTAGCA  
TTACTTAAAAINTGGTTCTTTAGAAAATAAATGCAGGTTATAAATGTGTGTATATTNAAGATTA

SEQ ID NO: 609 ACTTGTTACCAACTTTGCTTTCAAACCATTTTAGGTCCACAAAACTTCACTG  
AAGTGCTCCCGATCAAAATCTGATAGTTTCAACAAGGTATTTCTCAAATCGAGGTAAGTTGAGGTGC  
CCACTTTCATTAAATATAACCCCCAAGTTCTGGCNGGATGGTAACATATGTTCCATAAAGAAGAGGC  
AGTGTCATCATGATTAATATGTAAATGAGGTANATGAGGGATTAAATCATTCCAACAAGAAACNCC  
NTCAAAATCCAATCATCTATTATCCTTTCAATATCATATTTAAATGTGANCTTTTCTTTTAAATACTG  
AAAACTCTTAGTCAATATNCTCTCTCATTAAAGACAAGTGTAAGAANGTGATATGTT

SEQ ID NO: 610 ACGCGGGGACTACGATGGTGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTT  
CCCACCTTCTTTACGTTACAACCGAATGTGGACACTCGGCAGAAAGCAGCTGGCCGCCTGGTGCTC  
GCTGGTCTGTCTTCTGCGCCTGCACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGAGA  
GCCCCGCTCTTCAACAACGTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATTGTATTAG  
AGGAACTGAGGAAAAAAGGGAACCTCGAGTGGTTGGATAAGAACAAGTCCAGCTTCTGATCATG  
TGGCGGAGCCAGAAGAATGGGGGAAACTCATCTATCAGTGGTTTCAGGAGTGGCAGAACACTCCG  
TCTTTACCTGTATGAACTGACTAATGGGGAAAGACACAGAGGATGAGGAGTTCCACGGGCTGGAT  
GAACCACTCTACTGCGGCTCTGANGCCTACACANGAGCACAAGGCCNGATATACTGNAGCNTGCC  
GAGCGTAAGTTTACAGGACTGCTCTTACTTTTACTCCACTTTCAGGGTTAAAAGAAAAACCN  
NNCCCAAATGGTTTTGATCCAACATAAAGNCTCTC

SEQ ID NO: 611 ACATCATTGGGAATGGAGGGAAATAAATGACTGGATGGTCGCTGCTTTTTAA  
GTTTCAAATTGACATTCCAGACAAGCGGTGCCTGAGCCTGTGCCTGTCTTCAGATCTTCACAGCAC  
AGTTCCTGGGAAGGTGGAGCCACCAGCCTCTCCTTGAATAACTGGGAGATGAAACAGGAAGCTCT  
ATGACACACTTGATCGAATATGACAGACACCGAAAATCAGACTCAGCCCCCTCCAGCACCTCTA  
CCTGTTGCCCCCGGATCACAGCCGGAATGCAGCTGAAAGATTCCCTGGGGCCTGNTTCCAACCGC  
CACTGTGGACTCTGAGGCCTCTGCATTNCGGGTGGGCTGCCTGTGATATTTGTGATGGGCTGGTC  
TGGTCGGTTTTCCCATTTGTCTGGCCAGTCTCTATGTGTCTTATTCCTTTGGCTTCATTAAACANA

SEQ ID NO: 612 ACGCGGGGAAGAGGTGAAAATTTCCCTGGTAAATATTTCACTCCTTCCAAA  
AGACGCCCAGTTGAGTCTCAATACCTTGGATTGCAACTGGAACATGGAGACATCACTTTGAAAG  
GATACAATTTGTCCAAGTCAGCCTTGCTGAGATCATTTCTGATGAACTCACAGCATGCTAAAATAA  
AAAATCAAGCTATAATAACAGATGAAACAAATGACAGTTTGGTGGCTCCACAGGAAAAACAGGTT  
CATAAAAGCATCTTGCCCAAACAGCTTAGGAGTGTCTGAAAGATTGCAGAGGTTGACTTTTCTGCA  
GTGAGTGTAAGTGAATGGCATGACCAGGTGAGAATCCACCCCTGGCTTGAGAGCCCAAGATT  
TANAGTGAAACTCACACCCAAAAACCATAGGCGGAAATGTGACAAAGAAAAANCCCTCTCTGAT  
GTTCACTGGAAGCCAGACGAAAAAAAAAAAAAAAAAAAAAAAAAAGTCCCGGGGGAACCAANC  
TANACCCCAACCAACAACCTCTANAAACANTAAANNCCCCCTTTNCAAAATNGGNAATTTTGG

SEQ ID NO: 613 ACTTCAAATCACATAGCTTAAAAATATGGAGAAAAGACAGGTAAAAAAATAT  
CTTAACCTGTAGTAGTCTTTTTTCTTTTTTAAAAATTTTATAAAATACACTAATTTCCCAAAATAAG  
AATATTATGACACATTGGTGTCAACTTACACAAATGAAAGCTGATGCCAGCTTTTCTTAACATAA  
GTTCAATTCACCAAACTTTATATATTACATGAATATCCAAGTAAAGTATTTTTTTTTTTTAAATTC  
AACACACAGGAACATAATATACATTGTTTTTTTTTTTAAAAAAAGGTTAATTTAGGAATGATTTCA  
TCTTCACACATAAGGNTCTTTTAAAGACCATTGAACNNTTGCATTGCAAGACTTTCAGTCATGGTT  
CTTTTCCAGTTTTTATTCTTAANGAAAAAAGAGAAAGGTTTATTTTTTTTCCNGAAAAACCGG  
GGNAAGGACTATTNTTNCAGCTNGCTTCTTCTTGGTTCCGANT

SEQ ID NO: 614 ACGCGGGGGACCTTTGTAGCACCTCAACATGAAAGGGCATTAGCTATGTTTC  
CTGTTTTTACAGTGATCACCAAAACAGATCTTGCCACTTTGATTGTTAAAAATGAACCACATTCTAG

CCCTGGTCTGGGACTTTGGAGGGAGATGAATTTCTTGTTGGAAATGTAAATCTAGTGTCCATATTT  
 AATACTCTCACAGCTTTGTGTTTATTCTCTTTGCTCATGGAATAGCAGAACAGATAAACATGGGT  
 TAAGACATTTAGGAGAACCTGCTGTATCTAACCCAGTTGGATTTTCTTTCATGCTTAACACAGTAN  
 TGAATAATAGAAGGTAGGCCGGGCACAGTGGCTCATGCCTGTAATCCCACACTTTGGGAGCTGAGC  
 AGGTAGATCACCTGAGTTAGGAGTTAGACCCAGCTGCCAACATGGCGAACCCCTGCTTACCCAAAT  
 CAGAAAATTACTGGCGNTGTGGCAGCGCTGAATGCTNTCTTGGGACCCNNGCAAAAATTGTTGAA  
 NTGGAACAAGTT

SEQ ID NO: 615 ACGCGGGGGTTCACAGACCAAGCAAAGAGACGCAATCAGCAGCTGGTCTCAC  
 CCTTCTCTGAACAGTGACCAAACCTTTCACCTGATGAGCAACCTGGCTGGGGATGTTATCACAGC  
 TGCAGTGAAGTGCAGCTATCAAAGACCAGTTAGAGGGTGTGCAGCAAGCACTTTCTCAGGCTGCCC  
 CCATCCCAGAAAGAGGACACAGACACTGAAGAAGGTGATGACTTTGAACTACTTGACCAGTCAGAG  
 CTGGATCAAATTGAGAGTGAATTGGGACTTACACAAGACCAGGAAGCAGACAGCAAGCAAAAATA  
 AGAAGTCTTCAGGTTTNCCTTTCAAATCTGCTGGGAGGCCATTAATCTAGGATCANCTTGACAGAG  
 CACANAAACACCAAAAAAATTCAAACAGAAAAAAGGAAAAAGAAAAANTTGACTG  
 TACTTTATGATACTTAGATTTGTTTATTTCTCTGCANGAATTAATGNTTTAATCACTGNACTGTAA  
 TTGANTTTTGNCGTATTTTGGTANAAGCATGAANGACTTNA

SEQ ID NO: 616 ACCCAATGGGCAGGGAAGATCAGGAAGAGATCCATGGGACATAAGGAAGTT  
 AGGTTACTGCGCATAGCTCCAGGAGATTGTCCCCTCCATTTCTCCCACTCATTCTTTTGGCTGTGG  
 GAGTGTTCCTCTAAGCTGATCCACTACCAGCNGCTGGAACACTCCCCTAAAAAGCTAGCTTGGCT  
 TCCTGGGCAGCCAGTTCTTNAGGAGCAAGGCTTGTTAAGTTCTTGAAGCTGNCACTGCTGCAGGCC  
 CCCCCCTTANCGCACANAGGACAATCCTCACCAACACGCCACCANCCACTCNAATGGAATTCCTAT  
 TCCCAAGGCTGACGCCACCTTCTTTTCCACTTTGGGTGGTGAGGCTTGNAGTATCTGCCAATTACT  
 TCAATCCCTCATCTATTAATTCTGTCTCATAAGTGGTGATGGAAATCAANNNACCAAGTTAGGA  
 NGAANGACCCTNCCCAAATAGAGCCTTTGCAANTTCAT

SEQ ID NO: 617 ACAGGTCTGGCATGGTGGCCACCACGTGCATCTCCTGAATGATGTCATTTAGG  
 TCCAGCTCGGATTCCATGAACCTTCTCTGGATTGTCTGGAACTTAATCCGCAATCTTGTTTAT  
 ATGATCTCTTTTCAAATGTGAGGATCATTTTCTTCACTGAGCTTTCATCCAATGGCTCCTCCTCTTC  
 CTCTTCTCTTCCCATCTCTGTCAATAATCTGCANCAGCCTTTTTTGTGCATCATCCGCTTCCCTCCA  
 CCACAGTCATTTCTTCTTCCCGATAGCGGCGNGTTCTCGAGTACCTGCCGGGCGGCGCTCGAAAGG  
 GCG

SEQ ID NO: 618 ACTGCGTTTGGGCCTCAAAGGACATCCTTGAAGTCCAGTTTCACATCGTTGT  
 CAATATGAGGCATGGCGCTTANCCTCGGGGTAGCGATGAATNTGAGGGCTCGNGCAATCCCGCTA  
 AGCCTTCTGGGCCACGCCAACAGGTTCTCTNGTTGCCGGCAAGAACTACTAGCCGACGAATTCC  
 AGCTTNGGCGGCGGNAGGGCGAGGCCTTNTTGGATAACATAGTTTCACGCACTNNCCGANACTNTG  
 CTTTGTACNACTGTAAAGGGCA

SEQ ID NO: 619 ACTAGTAACAGGCAATTAACAACTAATAAGAAAAATCAGCATTTTAACAATT  
 TAAACGCTTCATGACAGGGTAATTCATGTCCAACATATCAAAAACATATTTATAGATAACTTTAGA  
 AAGAAAATACATACTTTTTTTTGATAATCACAAGTAGCAATGAGATTTTCTATATTATTTTCACTCT  
 CACTTTAGAAATGTTTTAATTGTCTAAATTTAATCAATTCATCGATTAAAGGAAAGACAATAAAAT  
 AGTAAAATTACATGTGTTTATATATAAGTGTGTGTGTTTCAAATAACAAAACGCAGGTTGTAACT  
 AAAATCACTGGAAGGCAAATTGAAGACAAAAGTGATGCTGGTTTAAAGTTGTTTGGTCTTCACATA  
 AATTCCAGTTCCAAGTTGTNTTCAACTCTAGTAATATCCAGATAGATAATTACCTGCACCTTTTCT  
 TTCTATCCTGTTTCTTGTCCACTGAATTTGCTTTTACTGAACTGTAA

SEQ ID NO: 620 ACTTTTTTTTTTTTTTTTTTTTTTGGTCTTANACTGGTTAATTTCAATTGCTTT  
 ATCTTTTCACTTTGCTGATTCTTTCTTCTGTCTGCTCAAATCTGCCATTGCATTCTCTAATGAATTTT  
 TCATTTTCACTGTTGTAATTTTCACTCCAGAATTTTCGATTGATTCTTTTTTACTATTCTTTATTG  
 ACGTTCCTTATTCATTACATATTCTCTGTTTCCATCAGACTTTTTCTTGGTTTCTTTTACCT  
 ACTGACTATATTTAAGAGAGTTAATTTGACATCTTNGGCTAGTAATTCAGTGCTGGGCTTCTTCA  
 GGAGTGGTTTCTGTCAACTNTTTTTATCTTGTGAAAGACCATACTTTTATGTTCTTGTGCTTTGA  
 ATTCTTTGTAAGACTGGATATTTTGCCTTGAATGTGGAACCTTTGGAATCAATTCCTCCTCTCAGT  
 GANGTGATTTTGTNTGAGGCTGCTGCTGCC

SEQ ID NO: 621 ACAGAACTCAGAGGAAAAAAGAAATTAATTTTAGCTTTCTGGAGAGCAGCC  
 CCTCTCTGGCACCATCAAACACTTCTTTGTTTCCCTTCAACTTGGAACCTTCAAACATCAGGGGTT  
 GTGAGGGTTTGGCCATTCTTTTATCTTGGGTCCATGTGAGTGACAGAAATGGTGCGGCTTGGGAAA  
 GATCTCCCTCCTTTACATTTTCTTCTCCTCCTCCTTATTCTAAAACTGTGCCTCCAACAGA

GGGGCAGGGGCTCTTGTAGAGAGATCCCTGGCCCANGACAGGAGATGCCAAATCTAATTTATCTC  
ACTGAGGGCCTTTGAGAAAAACGCTTCAGGGCCAGGCTCAGTGGCTCATGCCTATATAATCCCAN  
ACTTTTTTTTTTTT

SEQ ID NO: 622 ACGCGGGGGCCAACCTGAATGGCGTCTGTCTCTCATCCATATGGTGCCTGGA  
AATATTTACCAGTCAAGGTCAAGGTCAAGTCAGCATCTGTGGTTAAAAATATAGCATTCTGACCTAAAAA  
AGTTATTTTGCAGATGAATGTGTTTTCAACTCAGGACCTATCCAAATGAGGAATTTTTAAATATTCT  
TTTTTTTTCTATTTTAGACATCAATTCTATAGATTCTGACTTTTTCTAACCTCTTATANACATGCC  
AAATGCTGGCAAAAAGANGTGCTTTTTGGATATGGCAGCACTTGTAATAAATAAGCAGTNAGCAA  
AATCCTTTTAAACACAGAAATCCTGAGTTCCTCTCNTTGGTGGCTCAACAATTCGTANCAAAATAN  
ATCCTTTGAAAGAGAATTAN

SEQ ID NO: 623 ACGCGGGCTTTTACGTCGGCCTTCGCGAGCGTCTGGGCGGGTGGTAGGAACA  
ATGGCGCTGTCTTAAGTGACACAGTGGAGCANCTCTGAAGATGCAAAGATACACGAAAAAACTTC  
CNGAACATNTGGGAGAATATTTAATGGAAAATCGCTTGGTTAAAAACCTGACACTTTTAACAGTGA  
ACAGCGTTCTGAGTGTGGACGAGTAGCCANTGAAGATAATGAATGTCGAATGTGACTGACTAGCA  
ACTTCATTTTGAATGAGGGTCNCTNGTCTGCCATTTGATAGAGGCCAGATGNTTGAAAGTTCAAA  
NTTGCANCTATTNTTGCTAGTGCCANAAGGTTANTTGATGTGGGGGAAAGCTGTTAAGAAANCCN  
TCNANAAAAANNTCTTTNTTTACAACATGAAANAAAAA

SEQ ID NO: 624 ACTTTTTTTTTTTTTTTTTTTTTTTTNGAATTAAAAAATCCATTTTATTGCTTGG  
GTTTAAAAATAGTTGGGGGATACAAGTATTTACAATGCTATTGGAGTCAATTATTGGCAACACTTTG  
CAACAGTAATACCATTTCTAGCTTTTCAATTGGCAATACTTAAAAACCTTACTGTAGTGACCTGATTT  
TAAATACCATATTATATTTACTAAGTTAANAGCTAGTTTTTACTCTCTTCCATAATTTTATTACATG  
AATGTAANATGATGGCTCAAAAATGACGACTNATAGTTTGAATTTATGTGTATGCAATATACATAT  
GAGAACCAAATTCACAAGTGCATGATGGTACTCATGAACATTGATTGTATGGCCTGNCAGTTATT  
CCTTGGTCAATAANACTGAAGGGNCAACCCTTTCTTTCAAGAGTTGGCCTTCTCTCGATTTTAGA  
ATTAATTGGATATTTTCTCCTGCTCTCATATGATTAGNGGGAGGTTTCATCCACAAACAAAANACA  
GGAAANTTTTGCAACCTTNTCTGGANTTNCCTAGGGGACCAAAGGGATTNATTCAANANNNGNTTAA  
NNAAAAAAAAGCTCTTTTAAAGGAAAAACNA

SEQ ID NO: 625 ACCAAATGGAATGAATAGGGGAGAACATGCATTAGTTCTGTTTGAAAAGTGT  
GTGCAAGATAAATATTTGCAGCAGGAACATATCATAAAAAAGTTAATTAAGAGAAAATAAGAAGC  
ATCAGGAGCTCTTCGTAGACATTTGTTTCAGAAAAAGACAATTTAAGAGAAGAACTAAAGAAAAGA  
ACAGAACTGAGAAGCAGCNTATGAACACAATTAACAGTTAGAATCAAGAATAGAAGAACTTA  
ATAAAGAAGTTAAAGCTTCAGAGATAAACTAATAGCTCANACGTTACANTAAAAATGCCAGTTCA  
GCAGTTACACAANAGATGGCCCCACCGATGGAACAGGCCAACAGAATGTGAAGAGGGCCGCCA  
AGAAAAAGAAGCATGGTATGAAATATGTAAGAGGGGAGAAGGATCTTTAGATCTTCGAAGGAAAA  
AGAGCACTTGAGAAAAACCTTAGAGTGCAATAGGACTTGAGAAAAACACTACAATTAACACTTCTC  
AGAGAANGCGGTGCCCACTGTTGAACTAGGAGGCAACCGCTAGNTCTCAGAGAATGANTTTTAGG  
GNNNTTCTCNCGNCTTAANGAAGGGCCCAACATTAAGCNGGTGNTCCNGGACCAAATACNCTGA  
ACCC

SEQ ID NO: 626 ACAATGATTCTTAAAAAATCTTTGGCCTTAGTGGCCTTTTCTTCACTTACAC  
ATTA AAAAATGCTGCTGCAGTAACCAAGTGTGTTGGGAAAGGACATCAGTCTTCAAGAACCATAAACT  
GACAGAATTTCAATACAGTAGGTTTCCAAATTGCAATTTGTAGTGACATGACAGTAAGCGAGGTT  
TTGGGTAAATATAGATGAGGATGCCTATTCAGACAATCTACTTCAAGTAAAAAAAAAAAAAAAAA  
AAAAAAATTCACAGATACCCATCANTTCTACTTTAGGNTNTAACAGTGCTTAATCTGAGAAAAACA  
ATGCTTCACAAAATNTAAGTTACTGGGGTGATAATTA AAAAACCANGTGGTAATAACAATATCTCA  
AATCCAGGCCATAGCTGAATAATAACCAAGTCCTCGGNCGGACCCGCTAAGGGCG

SEQ ID NO: 627 ACAGAAATAAAATAATGGGAATTATCATTAACCTCACCTGGTTTTCTAGCTT  
AGTAGAACCAAACAGAAGAAATCATGGCAATAACCATTAACCTATAGAAAAAAGGTAATGGAAAA  
ATGGTTGCAGGTTTAAATCACAAAATGAACCTAATTTTTGTTGATTTTGTCTGCTAAAACACT  
AATATCTATAAATATGAACTGACAGCATCGTTCTAAATTTACTTCTGAAGAGCTGTGAGACTTCA  
ATAAAANATAAGCAAGTTACTGGATCATATTTATGGGCTGCTGAATTAACCTACCCGAAAAGTATC  
AGTTCTTTCAAAGAACACAAAACAAAGTGAACGTGGAAAAAAGCCTCTTTGCAAAGTCCTTTATTA  
GTCCTATCCTCTAAAAATCCAGCCAGANCTTGATATTCCTGGATTCTGTTTAAAGTACCTTAGTTTA  
AATATGACACTGGGATTGCACATGGGAAAGGNAGGATTGTGACCAAAATTTCTTTTCCAAGN  
AGCATTTCTTAATCTCTATCACTTTGCCCTNCCTGTCTGGCGGACCCTANGNGATCACCN TGNC  
GTCTGNGNCCACNGGNNACTGCGATATGCTCNGTCCG

SEQ ID NO: 628 ACGTTCCTTATATTCTTGCTGTCAATCAATCACAAATTTATATCAGATTAG  
GATAAACTAAGCCATTTTATGTATTTATTTTAAACCTTATTTTGGCAGAGTAATTCCTTAGAATTG  
GAAAAGCTGTTACTTTGAAATTTACCAATTTATTACAAAACATAGAAATGTATTGTAGCTACAAAGA  
CAACCAAGCATTTTCTGTGTTTTAATGAATATCTAAAAAACTACATTTAGTTTATTTTACTCAGTTT  
TGAATGATTTTCTGCTCTATTGCTTTANAAAACTAANAGATTAATGATTCTTTGGATAATTTT  
CTTTTCTTGATCTTATTTTACATTTTCGCAGAGNTATATTATAGTTTAGTAACAATGGCANNATGTNTN  
CNGGATAACNGAANCACTAANGTGTGGGNATTAGAATATANTNGTGAGCAGATGGACACTGG  
GNATATG

SEQ ID NO: 629 ACATGACCTTTAGTGAAGATTATTTGTTCATCAAATTACCCATATCCAAGTTTC  
CATGGGCCTGGAATTTCTTTCCACTTGATAGAAAGTATATATTAGGAAGTCCAGTTAATAGTATTT  
TTATTTAAAAAAGGAAAAAAGAATCAGCAGAGTCAAGTTGTCTTAGTCTTAAGGCTTTC  
TGGATTTCTTCTTGGAGGAGGTCAGGATCTTCCCAAGGCCTGGGTCTCGAATATTCTTCCAGTCT  
CAAACCTGGAGCTTTTGATTTTTCATATCCGACTCTAAAGATTTTATTCTCTTTCAGTTTTCAG  
CTCANGACCATTTTACTCTTTCAGCATCATATCGGNTTTGAGAACTCCGAAGACCAAAAGACCT  
CACAATCAGCANCAACATGGGGGACTCCTAGCCGAGAGCTTGTCTTGCNAAAGCACCATCACCAG  
GGTGCAACATGAGTGACTCTCCTCGCTANACTCCACGGCCTAGCCANACTCCAANCTACAGCTCN  
GCTCTNCAAACGATACCTGCCGGCGGCGTCGAAAGNGAATTACCACGNGGCGTNTATGTTNCCCT  
GTCAACTGCG

SEQ ID NO: 630 ACAGAGTTAACAAGTTTTGAGTTTTTATATAGGAAAAGCCTAGTCAATTCAG  
ATGCTTCTAGAAAAATTAACATTAAAAAACAAATAGAAATCCATGACTAAAGGGGGAAAAATAAC  
TTTCAAAGTTACCAAATTCGAATCATATCAGGGACCATTATAAATTTCAAACAGTAGATTTTACC  
ACACATATTGCATTTTCAAATTTCTAATGTAGCAAAACGTAACCACATAATTTGGCTACAGCTAATC  
GTTTCAGAAAAGNTTAAAAAATTAGCAAAGTTATATCTATAAACTTTTGGAGTTTCTTTTGTCAA  
GAAAAANGCTTAAATCTTAAATAAAGGAAACAAACAATCCTCTTAAATTTCTTATAAATAACTCTCC  
AGACATATATTACAAATCTGTGTAAGCTTTCTTCTGAGAGACTTCCAGGATCCTTATCCAAAGG  
ATACCTTAAAGAGTCTTCATCATTTCTCATGTGATATGATTAAACTCTATAAAGGGATGGGNATAT  
GCATCTTATCTGCCTTCCCATTTGTTCTTCTGAAGNATTNTTAAGAAAAAACTTTGTAAAGGCNNT  
NTTACCTCAAATGNTGCTTGGGANANTGTGAAANAAAAAAGAAG

SEQ ID NO: 631 ACTAAAGGAGATAAGTGTAAGTTCTCCCATGACTTGACTCTGGAGAGAAAAAT  
GTGAAAAGCGAAGTGTTTACATTGATGCAAGAGATGAAGAACTTGAAAAAGATACTATGGATAAT  
TGGGATGAGAAAAAGCTGGAAGAAGTAGTGAACAAGAAGCACGGTGAGGCGGAAAAAGAAAAAA  
CCAAAACTCAAATAGTGTGCAAGCATTTCTGGAAGCTATTGAAAACAACAAGTATGGCTGGTT  
TTGGGTATGCCCTGGAGGGGGTGATATTTGCATGTATCGTCATGCACTTCTCCTGGATTTGTGTTA  
AAAAAAGATAAAAAAGAAAGAGAAGAAGATGAAATTTTATTAGAAGATCTAATTGAGAGAGA  
GCGTTCTGCCTAGTCCAAATGTTACCAAATCCTCTAGATCTTTTCTTGCCTGGAAGAAAAGGAAAG  
ACAAGAAAAGATTGATAACTTGACAAGATATGGAAGAAGGAAGCTGCTTCAAAGCAGGGAAGCCT  
ATGACAGGGTTCGGAGTGTGATTCTCCTGACTGNCATGTGATATNAGAACAATGTCCCTCNCNG  
ACAGGGNGTGAGGTGTATCNGGTGAATGANTAATTTACCGCCTGCGGACCCTAGGNGATCA

SEQ ID NO: 632 ACGCGGGGATTTAATTTTTCTTTTTTTTAAAGTGGGGAGGAAGGGGAAGCTAG  
ATGGACTAGGAGAGACTTGATTTTGGTGCTAAAGTTCCCCAGTTCATATGTGACATCTTTTAAAA  
AAAAATAACAACAAAAAATGAGAGAAAAGCTAAAAAAGTAAGGGGNGAGCAGTTAATG  
GTATTCATTCCACATACAATATCTGNGTAAACNATTTTCTGGTAAAGTANCTTCNATGGTTTTT  
GCTTTTATAATACCGGTANGTCTATTCNTAAANCCTCTCGNCNATGCTTNTCTTNTGGGNTTTAA  
CTTTATATAACTTNAGA

SEQ ID NO: 633 ACCTTGGGTCTGAAGGGGTCACTCAAGCGCTCTGCTATCTCTTCAGCTAAAAAC  
GGGTGTCAGTTTTTCTGCTACTAAAGATAATGAGCATAAGCGTCACTGACCAANACTCCAGC  
CAGAAAGCTGCACATGTGACCGTGTCTGGGGGCACCCCAAAAGGCGAANGCTGTGCTTGGGACAC  
ACAANTTANNGACCATCACGGGAATTCTGCTGCTGTTATTACCCCATTCANNTNACAAGTGG  
CANGCANTANTCCANTCTCCAATAATAAACANTGTTTGATCTTAAAGNATGTGNGTNTCGTNNCC  
TCAACTATGACAANACAAANGAAANNTTANACCATGGGGGGCAATCCAAATGAAAANAATTATC  
NANATNA

SEQ ID NO: 634 ACTTTTTTTTTTTTTTTTTTTTTTCTTTTAAACATTTAATAGAACTNTATNCA  
ATAANTTTTTTACTATTTTTTACATAAGATAGCAACCACAGAANTTTACNTAGGTTAAAAAGCAAGAC  
GGNTAAGGAGGACCCANTCCTGTGGGCTGTTTTCTCAGAGGATAAAAAGCCAAAGTTCACCAGGG  
AAAGGGGTAAAGACTGCCCAATTTAAGTANAGGGGAANAAAAGCTAAACTGCNGGNTTTCANA  
TAAAGATAACCGATTTTAGGCCTTCANTTTNTCAAAGACCACTCAAAAGACAGTCCCCCTANTNC

TTAANGGGGATTCTNTCCNTGNTGGGGAAATCCCTTCAAANANNNNCTTTCTTTAGGNTTTGCCCA  
ANNAATCCAANAANCTTGATNGGATTATNTTTTTTTNNTTTCNTTGCCTNAGNGGAAGAAACCCNN  
GCCACNNCCNAAATTTGNTTTCGTTGACTNTTTCCCAACNAAAAAAAAACAAG

SEQ ID NO: 635 ACGGCAGGGAAGAACTGGAACTCAGAGAAAGAACTGCCCTTCCATCTAC  
AAAAGCTGAGTTTACTTCTCCTCCTTCTTTGTTCAAGACTGGGCTTCCACCGAGCAGGAGATTACC  
TGGGGCAATTGATGTTATCGGTCAGACTATAACTATCAGCCGAGTAGAAGGCAGGCGACGGGCAA  
ATGAGAACAGCAACATACAGGTCCTTTCTGAAAGATCTGCTACTGAAGTAGACAACAATTTTAGC  
AAACCACCTCCGCTTTTCCCTCCAGGAGCTCCTCCACTCACCTTCCACCTCCTCCATTTCTTCCAC  
CTCCTCCGACTGTCAGCACTGCTCCACCTCTGATTCCACCACCCGGGTTTNTCTCCTNCACCAAGG  
CGCTCCACCTCCATTCTCTTATACCAACAATAGAAAAGTGAGACATTCCTCTGGTTATTGAATAGNC  
CGTTCTGCACGTGCATTTCAATATGGCAAAGTTGCCTTTCCCCATCTTCTGGGTCTGCTCCTTCNTG  
GGCCTAATTCTTGNGGACACCNNGCCAAGCCNNTGGGGANCTATTTATNGCNTAGGAAAGAGAG

SEQ ID NO: 636 ACGCGGGGGGGTTCCTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCG  
TGGCAGTATCGCTTCCACCCTTCTTTACGTTACAACCGAATGTGGACACTCGGCAGAAAGCAGCTG  
GCCGCCTGGTGCTCGCTGGTCTGTCCTTCTGCCGCTGCACAAACAGTCCAGCATGACGGTGATG  
GAAGCTCAGGAGAGCCCGCTTCAACAACGTCAGCTACAGCGAAAGCTTCTGTGGAGTCGAT  
CCAGATTGTTTTAGAGGAACTGAGGAAGAAAGGGAACCTCGAGTGGTTGGATAAGAGCAAGTCCA  
GCTTCCTGATCATGTGGCGGAGGCCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTTCCAGG  
AGTGGCCAGAACAACTCCGCTTTACCCTGTATGAACTGACTAATGGGGAAGACACAGAGGATGA  
GGAGTTCACGGGCTGGATGAAACCACTCTACTGCGGGCTCTGCAGGCCCTACAGCAGGAGCACA  
AGGCCGAGATCATCACTGTCAGCGAATGGGCCGAAGGCGTTCAAATTCIT

SEQ ID NO: 637 ACTTTTTTTTTTTTTTTTTTTTTTTTNGAGTTTTAGGTATTTATTTACAAAGTTCTTA  
CTAATACAATTGCTTTTAAAATGTAGCAAAGAGTCATTTACTACTCTCANAAGTGGCACATACATG  
GCATANAAAACAATCTATAGTCAGTTAACTATTAACAGAACTTGAAATTTAAGTGACAAACA  
TTTGTAGCACTCCCTAAAGAAATAGGAAATAAAAATGCATTTATCCATATGAACTTGATTATTCTG  
AATTACTGACTATAAAAAGGCTATTGNGAAAGATATCACACTTTGAAACAGCAAATGAATTTTCA  
ATTTTACATTTAATTATAAGACCACAATAAAAAGTTGAACATGCGCATATCTATGCATTTACAGA  
AGATTAGTAAAACCTGATGGCAACTTCAGAAATTTATTTTCATGANGGGTACATTTTGATAGTATTTGA  
TAGGCTTTTTTCCAGGTCAAATTAATTTAGTTGCTTGCNAAATATAAAATCAAGCTTGCTCCAGTTC  
CACAAGGACTCCNCCACAGTCTTTAGGATGGGANAAAAAATCACTNG

SEQ ID NO: 638 ACNTNNTTTTTTTTTTTTTTTTTTTTTTTTACATTCTCANCCCATGCTGCTTTATT  
GCTAAAATATATNTNTACNCAGGGGCAGAGCTTCCAACCTTTTAAACAGNTCCTTTTNCATCAAA  
AAATTAATTTNTTCAAAAACCCCNAAATGCATANATTA AAAAGCAAGCTGCCTTTTGTAATACT  
TAAAATAATATTCAATGAAACTCTTAACGTTNTACNCCAATGGNGCAATGGANAATATGGAGGAC  
AGCAAGTAAACNCAGTGAGCAACTNTTATCTTACAAGGCAGTAGGTAAAGTATAATGTGAAGGA  
CAGNCTAAGATAATCTTTCTGGTTATAAAAATGGGTGCGTTTTGTGATAAGTGCCAAACTGTTA  
CTGTTTAAGGATTGATGATAATGTATTACCAGAATACCTTTCTTATCCCTGGGAGAAAAATTTTTTT  
CACTGATGGAATGACATATCACCAAGGGAGAAAAAACCTGGATACTGNCAGTAACCTAANATACC  
TATTCTGACTTTAAANAAAAATTAAGGACTTCATGGGAGAAAAAT

SEQ ID NO: 639 ACAGAAGGTAGATTTAGAGGTGAGGGGCAATAAGTAAATGACTCGCACAGA  
TGGTGTTTCATGCTCAATGTTTCTCTTCTGAACAATTTTCAGATTTGGTGTTAAAGCACCTTCTGGCC  
AGGCGTGGTGGTTTATGCCTGTAATCCCACTCTCGAGAGGCGGAGTCGGGTGGGGATTGCTTGA  
GGCCAGGAGTTCAAGACTACCCTAGGCAACATAGTGAGATGCTCGTCTCTATTA AAAAAAAAAAAG  
AAAAAAAAAAGGGGACCCAAGGCCATATATCTGTTTCATTCTAGAACTCTCCACAGTGTGAGAT  
GAAGACTGCAAATCCAGTGGTCTACATTCACAGATAAATTCCTAGCCTTTGGTTAGTGAGGAGAAT  
GGGATGGGAGGACACAAAATGAGAACTTTTTTGTGGTATGCTCAATTCGTATCTGGCAGTACCTCG  
GCCCCGACCCCCCTAANGGCG

SEQ ID NO: 640 ACGCGGGGACCTACCTGGGATAACGGCGGCGAGCGGACGGCTGCATTTACG  
GGGTCTCCCGAGGGCCAGAGTCGTGGCTTACAGAAGAGACGAAATGTGGTCTGAGGGACGATAT  
GAATATGAAAGAATTCCGAGAGAACGAGCACTCCTCGAAGTCATCCAGTGATGGCTACAATAG  
ACTAGTTAATATTGTGCCAAAGAAACCACCACTGCTANACAGACCTGGTGAAGGAAGCTACAATA  
GATATTACAGTCATGTTGATTACCGAGACTATGACGAGGGCCGAGTTTTTCTCATGATCCNAGAA  
GTGGCCACCTCACAGAGGAGATGAATCTGGTTATAGATGGACAANAGACGATCATTTCTGCAAGCA  
GGCAACCTGAATACAGGGACATGAGAGATGGCTTTANAAAAAAAAGTTTCTACTCTTCCATTATG  
CGAGAGAGCGGCTCTTATAAAAAGGGACAATACTTTTTTTCAGAGAATCACCTGTTGNCCGAAAGGA

TCTCCACAGCAGACTGGNTCCTTGTCAGTAGCAGANCTCTCTCCAGAAGGGCAGATCATCTCT

SEQ ID NO: 641 GGACTTTTTTTTTTTTTTTTTTTTCAACTTAATTTACTTTTTATTAACCAACC  
TGCTGAAAAGAATACAAATTTTGACTAGGAGGANAAGGGGGAGNGGATTTATGATTTTATGGCA  
ATTATAACAGGAGTGTTCATGGAGACTTTAGGCAACTAGAACTCANCATGAAACCANAAGGTA  
NAAGCTGATAAAGGCAGGGGTAAAAAGGAGAAAATGAGTTCAAAACAATACTTTTCGGCCAGG  
TGTGGTGGCTCACACCTGTAATCCCAGCAATTTGTGAGGCCGAGGTGGGCCGNTCACAAGGTCA  
GGANATGGAGACCATCCTGCTAACGTGGNGAAACCCCATCTCTANTAAATACAAAAAATTNGCCG  
GGTGTGACAGCATGTGCCTGTNGTCCCTCTACTCAGGATGCTGAGGCAGGAGAATGGTGTGAACC  
CGGGANGCAGACTTGCAATTGAGCAAAANATCGGCCCNTGCCTCCAACCTGGGCCGNACAGACGAGA  
NTTTTTTNAAAAACAAACAAAACACTAGAGGGCATTGTGCATGCAAATTNTGCTAGCATATT

SEQ ID NO: 642 ACTTGAGTGCTGCTGAGCTTCAGCCCACTGAGAGTTTACCTCTGGAGTTTCACT  
GATGACTTGGATGTTGTGGGTGATGGTATGCAGTGTCTCCTTCACCTTTGCTTTTGTATCCTTCAC  
TAACCTTGAAGATCATTTAGTCAAAGAAATTGCTGAAGACCCAGTGGATATTTGGGCCAGATGC  
AGATGGCTGGAGATGGGTGCAGATCCCAGGGATCTCGAAATTTGAGAAAGCCTCTGCACCATTG  
TCTCAGAGTGGATTGGCTACTGCAAATGGGAAGCCAGAACCCACTTCTATTAGTTGATAGTTTGGG  
GAACCATTTTACTTTGGTGGATTAAATTTCTCGTTCTTCAAAGAAATTTNTGATCATCTCTCAC  
AAAACACAGGCAACCCAGACTACCTTGTTTTCTTCTGGATTGAGTATTATAGTCAAACGTAAAGC  
CAAATTTGNGGACGTGGACCGTGGGTCTTCCCACTCAGGGCAAGGAAGAAAGANAGATATC  
CCTATCAAAGCTCTGGGATTATTTGGGACTATACCTAAAGGAAGTGACATCT

SEQ ID NO: 643 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTNCACTTATTATTTATTTAT  
CTTCCAATTTCTCTTGCCAGACTCCCATCCAAAGAGTCATAGCAGCCTTNTTCCACCTTNTACATG  
AAATACATCCCCACCTGAACAAAGGCNCACGACAGGAGGGGAATAGGACTTNGCAAACCTGG  
ACACGGNATCGTTCAAATCTGGACTANGTCCGTTGTTACTGGTTTCACAGTTACAGGCTTCGGAT  
GGTCTGCNCGTGCTGTTTCAANACTAATGGNAGACTCTATTGCTTCTGTTATGTCCTTATNCAACCT  
GGTCAGCCTGCCTCTGCTCAAATATNGNGTAATCAANTGGNGAAATCTGCCTAAANNCAATCATAA  
CTGGGGGNGACTGTTAATAATANACCACCTGAAATATTCACTCTCCAGNCTTTNTTCATCCTCA  
TATTCTTNGNCAAGATAAGTGGCACAGCAAAAATGGNTTCAAAGAGGAATCCATTCTGGATTTTG  
CCTTTNTGGCCCCNCGTCTNNGCCGGACACNCTAAGGGCG

SEQ ID NO: 644 ACTCTAATTTCACTAACTGCCAAAAGGTTTTCCAGAATAATCTCAGTTGCTTC  
ATTCCTTTAAAGATGAAGCCCAAAGACGCATGGCGATTACTTTAGAGGAACAATTAGCAGCAGA  
GGCAGGGCTGTGCTGATCCCATCTGGCATCGCTGGGAGCTAACATTAAAGACATGGCACTTTGGG  
TCCGGGTCCAGGTCCTGGTTCAGAGCAGCTGCCACACCGTGGCTACTAGAGGATCCTTTTCCGGCT  
TTGGAAACTGAGGCTGACTGCACCATCATCTAAAGGCCTGAGACTGCTCGCCGTGCTCAACAC  
CGACTGGAGTGGCCATGTCTTCCAGCCACGCCGGCCGACCTCGGCCGCGACACGCTAAGGGCG

SEQ ID NO: 645 ACCCTAACCTGACAGGAATTAATACTACTGTTTTTTGTGGGGCAGAAAGCAAA  
ACCTGGTGTGTGACTTTTATCCTAATGGTCTTAGGCAAGGTTAGTGAGAAGAAACACAAACCCA  
NATGCATGCATTGTGCATTATTTGTAGACAAGCTACTTTTTCTTCTGNCCCTTAAACAAATTTGCA  
GCAATTACCTCCCTTTGGGGTCTAGAGTGAAAGCTAATTTGTGGGTAGATGAGATTGCANAAGA  
ATGGATGTCCATGGCTGTGAACACTGCACACTGAACATTCATCTCCAGTGCTCACACTGTGCAGCT  
ACCCTCCCTGGCTGCGTGCCATGCTGTGGGTNAGATTGTCACACATAAATTCCTCAGGANGAG  
TTTGCATGAGCATCCCTCGCAATATTCTGTACCGCAGGGGAAAGATGAAAAATTTACCAAGCATA  
ATATAGCAAGGACTNCCCCTATCCTTCTGCATAATGAATTAATACTAGAAATAACTTTGCAGGAGAAC  
NAAGCTAAGACCCCCGAACCANACGAGCTCCTAANACAG

SEQ ID NO: 646 CTTTTGTTNGCGCCGAGGTACACAAGTAAAATACTACAGAAATTAATTTCTT  
TCAGCATTGAAGTGTTTGCTTCTCTTATTTCAACTAAGTTGTAATAATTTCTGGTATCCTGGTA  
ATTGATAACTAAGATAATATCTACAGAGNGGATGCCATTAATTCTCTTAGCAATCACGTGCAGAA  
GGGAAGGGTTTGGGCCAACCTAGGGNTGTGCTGNCTCTGNCGTAGTCTAGGAAAGCACCCAGGN  
GTCAGGNATTTGGTCCCTTCCACCACACCTCAGCAAATGTAAAACAAGCCCACAGTTAATGGCCT  
GTGGCAACCCCTGGCCTGGAGAATCTNANNANAGCANTTGAAATGAGGAANCAANATTACATT  
TAAATATCTNGCTTTGGCACACAGTGGCTTTTATTGATTAATAAGCAAAT

SEQ ID NO: 647 ACCACTCTACCCTATACTCCAGGACTTCATCTTTCTTACTGAAAAACCCAAAC  
CATCAAATGTCCCTTGTCTACCCAAGAATTGTTCTGCAACATACGCCACCTACTTTCTTGCTTTC  
TATGAGGCTGAAGACTANACAGGCAGGAGTTTATTATACAGACCTGCCATATACTTGAAGTAGTG  
TGTTAAGTCAACCAGAGGCCAAGTTCTTGAATGCTTTCTCTGTTTGGATTATTTCCACGCTGCTGGC  
TGCTGCCTGGACAGGTCTGTCTGCACATTCTCAAGGNTGAGAAACCCCACTAGGGNAACTGC

CCATCAGAAGGNGAGTTCACTGGGTAATGGGAGCAACTGCTCATTGTGAAGGTGACATGATGTTA  
AAAGGCTTGAAGGGGCTGCATGGTGGCTAACGCCTGTAATCCCAGCACTTTGGGATGCTGAGCA  
GACGGATCACCTGAGTCACGAGTTCGAACTAGCCTGCCAACAAAGGAAAACCTCCGCTNTCTAAAT  
CAAAAAAAAAAAAAAAAAAATTAGTACCTNGGCCGCGACACNCTAAGGNG

SEQ ID NO: 648 ACTGGATTCTATAGGTCCACCATTAAAAGCTGGTATGGGACACCAATTTATAC  
ACATAAGGGTTAGATTAAAATTTTAAATTTTTTGGTTGATATTAACTGAAATTTATATACTGAGG  
TCTGAATCTTAAAAAAGTAAATGATAAAATTTAATATAGTTAACTGTTCACTGATATGTCTATT  
CACTTCATCATAACCTATATATTTAATTAAAAATCAAATTATGAGTCTGCAAATCAGATGCTATCA  
AGCAAATTGCCATCCAGGGTCCATAATTTCTTTTATATTTTATCTCAGATGAATATATACGATTCA  
GTAATTTTAAATGTTCCAAATGTTCTAAAAAATAATTTTCAAAAGCTTCCAGTTAACAGTTGG  
CTAATTCATTTGCTCCCAACGAACTACCTGTTTGTGTTGTGAGGTAGCATCAAAGACTATTGATCTT  
CTGTGACAGTAGTAGCCTTAATTCATACGCATTCCCTCTTCATAGGAAGAGTATGGACAACCAAAA  
AAGGGACNNATGANGTCNCCCTTTCATTAATCATTTGCCTCCNGG

SEQ ID NO: 649 ACCGGAAGAAGCAGCTGGCAAAGCAGCTCCCTGCACATGACCAGGACCCTTC  
AAAGTGCCATGAGTTGTCTCCAGAGAGGTGAAGGAGATGGAGCAGTTTGTGAAGAAATATAAGA  
GCGAAGCTCTGGGAGTAGGAGATGTCAAACCTCCCTGTGAGATGGATGCCCAAGGCCCAACAA  
ATGAACATTCCTGGAGGGGATAGAAGCACCCAGCAGCAGTGGGGGCCATGGAGGACAAATCTG  
CTGAGCACAAAAGAACTCAATATTCCTGCTATTGCTGTAAACTGAGTATGAAAGAAGGTGACCCA  
GCCATCTATGCCGAAAGGGCTGGCTATGATAAACTGTGGCACCCAGCTTGTTTGTCTGCAGCACC  
TGCCATGAACTCCTGGTTGACATGATTTATTTTGAAGAATGAGAACTATACTGNGGGCAGACA  
TTACTGTGACAGCGAGAAACCCGATGTGCTGGCTTGTGACCANCTGATTTTTCANCAATTGAGTA  
TACCCAGGCAGAAAACCCAGAATTGGCACCTGAAACACTTTTGTCTGCCTTGA

SEQ ID NO: 650 ACTTTTTTTTTTTTTTTTTTTTTTTTTTCGANAAAAAACCGCCAGTNTTT  
TATTNTCATGGAAAAACANAACAAACCCNCAAGTTGGAGTCNCGGANATAAAATNCANATGAAA  
TGAAAAACGGTCTGTTGTCATGAACNTCACTTTCAAATACCATTTTATNTGGAAGTTACTTTACTG  
NNGGGCAAACANAAGGCCATGCTGGAGTCTNTTACTTTTGGAAAATGGANAATCAAAAATTTGCT  
ANTCAACAAACAAAAAAGNGGGAACTCCTTGGTAAAGCTNTACAAACATAATTATNCNTTNT  
NTTTTACCAATAAAANATAGCTAGGGTANAAAAACANATGGTTANAACTGGNGCCAAACCAA  
AGNGAAAGCTTTGGTGCCTTCTNTAACTCCTATCCTGTTTCTTTAAAAAACACCAGTTTCTAAA  
NAAAGATTACTCTGAATCCCCAGGGTCTTTNCCCCAATTCATCCCTCCCTTTCCCCCCCCAAAAC  
CCGAAAGGGGCCTTGTNTTTTTTNTCCGGCNCCACTTTGGGAAGG

SEQ ID NO: 651 ACTGATATAATCTAACAAATGAAGGTGCACCTTTACTTCTGGAACATAGAC  
AGCCACCTTGTATGGTTGGGGTCCAGGTGATAATACAAAATCATTAATTTTTTGAAATGCAATTT  
ATTTGCAATTGTGTTAAATTTGTTGTTTTCAAAGAAGTGAACTTCATTGTTAACATTGCGGGCACA  
AAGAGTTTCATCTTCTGACCAGGATGGACACCAATTTGCATTTTTTCTGGATGAAAGATTTCAA  
ACATGTCCAGTTTTCACATCATAAAGTTGTAGGTTGGGTATCCCAGCTGTGCCATCTTTAGAAGT  
AGTGTAAGGCTGCCACGTTGCCAGGACAGTATTTTGGGTGAGAATTCAAGGCAAACCTGCCTTCA  
NGAGGTGGAAGGAGTGAGTAGTCCCTTGTTAGTGACACTGATAATTTTACTTTTCTCCATTGC  
CCCAGGCAAACAAGGTCCCATCCTTACTAAAAGATACAGACTTTGCAATTTCTCCAGATTCCCTT  
GGAACACTGTGCTTTCTGTAATGTGGTGGTCCATTCACCATGTCC

SEQ ID NO: 652 ACTGCCAAGGACAAGTTGATTTCTGGCCAGGCAAAGTTAACTCAGTTTTTTAG  
ACTATAAATTTGTGTCTTATATGCTTTAGGTTTATGTATCTATAAACCAATTCACCAAAGACATGCTT  
AATTTTTAAGAGATCAAGGTGTAAATTATGATGATTTATTATTTTGGTCTACAGTGTATGTAAGGTT  
AGTATGTTAAGCATTGTTTAAAAATACTAGTAAGTCATAATTATGCAGAATTTTACAAAGTTTAA  
TGCACAGAGAAAGCATATCATTTCACTGATACATCTTAACACTACTTTCTTTTAAACAGAC  
ATTTAACATACACAAGTTT

SEQ ID NO: 653 ACTTTTTTTTTTTTTTTTTTTTTTTTNGGCTTTCAATCTTTTATTTAAATGCCATG  
ATCCAGGATGGATTTTANATCTGTGTAAGCAGCCACATCCATGGACTGCACATAGTCCTCAAAA  
GCAGNGATCTGCTCCTCCAGCATATCTGTTCCAACCTTTATCATCTTCAACTACACACTGTATTTGAA  
GTTTCTTAATTCCGTATCCCACTGGAAGTGTANATGAGCCCCANACTAAGCCGTCTGCTTGAA  
TGCTTCTGACGCACTCCTNTAATTCGCCATATCTGTCTCATCATCCCAAGGTTTACATNTAGTAA  
GATGGAAGACTTGGAACAAGNGCAGGTTTTTGGCTTTCTTTGATTCAATTGNGCAAGACGTTT  
TTCCCTTAGCCTCTTTGCTTNTTCACTTTCTCCTCATCATCAGATCCAAAGAGGTCAATGNCATCA  
TCATCTTTACTATCNGNAGCTCCACTTCTGTAGNGTNTTCCACATCGGCNGGACCATTTTGCCC  
AAAGCTTTCTTCACTCCTGGCAGGCTGGGCCTTTTCTTTTCTGTAAGAACTTGATGNGGATTTTATACC  
AAACGGTAGGGCATTGAACACAAAATTCGGNCTNGCNGGTGGGGCCTGGGACAACCGGGGTTTA



AAATACTGGCNCCATTTGNNTTGNNGGATGGNNCAAACCCCNNGATGTTAACCTCTT

SEQ ID NO: 654 ACTCTAATTTCACTAACTGCCAAAAGGTTTTCCAGAATAATCTCAGTTGCTTC  
ATTCCTTTAAAGATGAAGCCCAAAGAACGCATGGCGATTACTTTAGAGGAACAATTAGCAGCAGA  
GGCAGGGCTGTGCTGATCCCATCTGGCATCGCTGGGAGCTAACATTAAAGACATGGCACTTTGGG  
TCCGGGTCCAGGTCCTGTTTCAGAGCAGCTGCCACACCGTGGCTACTAGAGGATCCTTTTCCGGCT  
TTGGAACTGAGGCTGACTGCACCATTTACAGAAATTTTCATCATGAAGATCCTTACACTTAAACACC  
TTAAGGTGATGAGTTTTACATGACTGTTTTCAAGGCCACAGCCACCACCTTTAATCATCATCACT  
AAAGGCCTGAGACTGCTCGCCGTGCTCAACACCGACTGGAGTGGCCATGTCTTCCAGCCACGCGG  
CCGACC

SEQ ID NO: 655 ACCCTAAGGCAGGCCCACTGGCTCTTTTTGATCAAGGATTCTGAGAAAAGCT  
GCCCTTGGAGGCCCTTGAAATAACATAGGGAGCAGAATGAGTGTCTGAGTGTGCTGACACAGT  
CCAGCTCACACTGCCATCACAGAGGCTGAGTGAGCAGTCAACCCAGGGAGGGGCTCCCAGTCTAT  
TCCATTCCCATGGGGCAAGTGACTAGAAGGTAAGAGCACCCGAGTAAGCCAGTGCCTTCTGTAT  
CCACACCCAGGAATACCTTCCAGTTGTCCAAGCAGCCGTAGTTGTAAGGATTCCTAAATACTCTGC  
CCTTGGCCTGTAGCCGACGTCTCTCCTTCTTGTGATGTGCCTTTCGATGCTAGTCTCACCTCGACT  
GATGAGAACAGCATGCCATACAGTTAGGGCACCCAGGGCAAGTGCCACAGAACTGCACAGGAAC  
CAGAGGTAGACAAGACTCTTGTGAGTCATCCTTTCTCGAAAGGAGAANGTGGGTGGTGGGGTCTT  
GGTGATAAGTCTGGGTGGCAACCCGCCTGTAGTTTGGTCTTGTGAGCTGTTTTCATTTTTCTCATGG  
CAGCATTAAAGCCTCCCGAAAANGGCCAACTTCCATACTGCAGTTANACACAACCCCANAGTCT  
TGAAAAAGCCNNAAAGAGAAAGAAAGTTCCCTTGGCCGNGACCACCNCTAANGGGCGAATTTCCA  
CAACACTTGGCGGGCCGGTTCTTATGGGATTCCNANNCTCGGNACCAAAC

SEQ ID NO: 656 ACACTTCAAATGCTGATTCTGTTGAATCTTCTTATTTTTTTTCCCCCAATGAAC  
TTAGGGCCTATGATAGAGAAGATACTGAAACATGCTAAGATAATAATTGGTCTTAGAAAGACTTT  
ACAGCAAGTATTTAATGAAAGCCACTACACTTCCAGCTCTAAAAAGAGCCAGACACAATCCCTGC  
TTCCAGAGTTCACCATTCAGAATACAGGAGGTCAAGCACTGATTCTGAATCATTAATGGGCTTTGA  
TTTCGTCTTTCAAGATTTGCATAGATACCTGATTTCTTCCGGAAGGGCAAATAAAATCTTCCCTTACA  
CACATCTTCTTTGGATTCTCATAAGGGTATTTTGAAGTCTGTATGTTGACAAGGCCAGTAGATA  
CTTTGTTTTCTCTAACTGCCTTCTTCTACTTAAGACTATTTGTTCTCAAGTATTTTGTCTTCCCTT  
TCCCCAACATACACCTTCTGCAGTTTTGTCTATCCAGTCCGCTTTTCTGTAACACTTTGCTCCCCATT  
CTGTTCCTTCACTGTTTTCTCTGTTGCTTTCCTCCTTAATAGAAGCCTGGTGATCTGGGCTTCCCAT  
CTTTCTTTCCTGGCTCCTTATTCATGGGTGACTGATCTTAGCACTTCCTTAAAGTTGGGGTTTTTC  
AATAAATATAACTCATCAGCCTTCCGATCTTGTATTTCTTGGACCTTTTTTTTTTGG

SEQ ID NO: 657 ACTGTAGAATGTGATGGAAAAGCATTGATGAGAATTTATTGGCAGTTCAGAT  
TGTGTTTTCCCACTTAGGCTCTTTATTAATTGGCTAAGGTTTTCTCCAAAAAGGGCNTTCAACNA  
TGGGAATTATTTNNTGTAACACNNGGCACATATTACCTATCTTCTTCTCTGCTTTGTGNTCCACC  
NNCNCTTNCNNCTCANNTCCACATTTT

SEQ ID NO: 658 ACTGCTTTTATTTGAGTTTATGAACAGAAATAGAAAGTATGGTGCTTGGGTTT  
TGCCCTTTCTTACTCCTGAAAGTTAAATCAGAAGACACTGATTTTCATTTTGTGAAATTTAGCTCAGA  
GACTATTGATCTTTTGTTCATTAATATGAACAACTATTAGTAAAAAATAGCTTTAACAGCATTCT  
GCTGATATCTAGTAATCTATTCTTTAATGTGAAAATAAGATAAAATGTCCTGGAGCTAATTCTAG  
CTTAAATTTGCCAGTATTTCTGTATGTCATTAAGTTTTTTTCTCTAAGGTTGGTAATAGAATTTGT  
TAATCTTTGCATACCTGATGGCATCTATGTCAATGCTGATTGGGTAATTATAAATCTGTGCTAATT  
TAAACTTTAATTTGCCTCTTAAGGTGATTGCTCTGAGTAATGATTGTAGTTAAATGAAGTATAG  
CTTGCAACTATACTATCACATGGGTGCTTAAGTAAAAATAAATAAACCAAAATTTGTCTGAGACAG  
GCTAAGATCAATCTTCTCATCAAACCAATTTTCTCTAAGAGCAAATTTCTTTTCAAGTTTGGGNG  
GCCATTCTTGAATGCCTCAAAATTAACCGTTATCTATTTAAATCTTCCCGGAATTAGNCTGGGAC  
CAAAAAGGANGGGNGNGGATNTNTTAAANGNGTAAATATTATCNCCATATTTGGGGGNGG

SEQ ID NO: 659 ACTTCTATATATAAATTTGGACGAATAGAAGTAAATATGTTTATTGGTGAAAA  
AGAATTCAGAACTAATGGCAGATCCTGGAAATCCAGACTTGTATCATGTATTAAGTGTTATCTG  
GCAATTAGCTTGTGAGATTAAGGTTCTGCACATGGAGCCTTGGTCATCATTTGATATATACACCCG  
GAAAGGGCCGCTGGAAAACCCAAAGCGTAGGGAATTATTAGACCAATTACAACAAAAGCTGTATC  
TTATTCAAATGATTCCTCGTCAAAATTTATTTACCAAGAACTTAACACCTATGAACATAATATATT  
TTTTCACTTGTTAAAGCACTGTTTTGGGAGGCGCAGCGCCACTGTAATAGACCACTTACGTTTCATT  
GACTCCACTTGTATGCGAGAGATATATTGATGCAAAATAGGAAAACAGGAGGATGAGAAAGTAGTTA  
ACATGCACCCTCAAGACTTCAAAACACTTTTTGAAACTATAGAGCGTTCCAAAGATTGTGCTTATA  
AATGGCTGTATGATGAAACCCTGGAAGATAGGTAGCAACTAGACTGTCGTTTTTGGTGAGCGGT



TCATTTNTTTGGAAACCTATGACATGAAAACCAAATTTTGAAAACCTCCATTCTTTTANCANAANGN  
TACTGGTCTTGTCTTGCCCAANCCAGCCAGATCANTTTTTCTTAAGCCTGATNTCATTGGC

SEQ ID NO: 660 AGCAAACCTTTCTGGATGCCCAACATGATTTTCAGTAACCACCCTTTAGAGTAT  
TTGTTTACTAAGTTCACCACATTTTGAACATGGTAGTTTTAGACTGCAATAATTTTAGACTTACAT  
TATACTTACTGCTAAGNAAAATCTAAATCCTGCAAATGCACAGAATTCAAGCTGAAATATAATGA  
TTTATGTTTAGCTCACATTGAAGTATTGGTTGGTTACTTATGTATTAATGCAGNGTGCATTACATT  
TAATCAGGTTTAGTCTGTTTCTATTTTAATAATTTTAAAAAATTATACAAGCANNATTAGATATTAGA  
CATGNTAGTTACAATGGNACACATTTTAGGNGTCGAAACACAATTTCAAAAATCCCTAATGNAA  
GGTATAAAAAATGNACANGAATTGAAAAATGGCCAAAGTANCNAATATTTTTCNAAGCNAATTTT  
TTTAGNAGGNATAATTTACATTTTGCTTTTCTAGNGGGGTGAAATGTTTATNGAGATTGGCTTGT  
ATTTATANTTTTATTC

SEQ ID NO: 661 ACTGGATTCTATAGTCCACCATTAAAAGCTGGTATGGGACACCAATTTATAC  
ACATAAGGGTTAGATTAAAATTTTAATTTTTTGGNTGATATTAAACTGAAATTTATATAACTGAGG  
CTGAATCTTAAAAAAGTAAATGATAAAAATTTAATATAGTTAACTGTTCACTGATATGTCTATTC  
ACTTCATCATAACCTATATATTTAATTAATAAATCAAATTATGAGTCTGCAAATCAGATGCTANCAA  
GCAAATTGCCATCCAGGGTCCATAATTCCTTTTATATTTTATCTCAGATGAATATATNCGANTCAG  
TAAAATTTTAATGTTCCAAAATGGTCTTAAAAAATAATTCAAAAGCTTNCANGTAACAAGNNG  
GCTAATTCATTTGCCCCANCGAATACCTGGTTTGGGGTGGTGAAGGNANNATNAAAANAANTAT  
GAATCTTCTGGGACANTAGT

SEQ ID NO: 662 ACTTTTTTTTTTTTAAAGATTACTAAACATACAGGAAGTGATAAGAAGTATCAT  
TCATCAGAAGCATCATTCATCAATCAACTTGAANAAAAAGGNGATATATTATTTCTTTAAGGTGCT  
GNGGATGTGTTAAGAGCATATTAGAAGGAATGGTTTTGTCTAATTTCTTCATGAGTTATGGNNGC  
TGAGACATCGAGTCTATATTTGGGGCAAAAACCTAAACGGNAGNACNAAAGGAAATCTATTNTAA  
TAGNATATTTTTGTTGAACANAGGAGGTTNGATAAGAACTGCAAACCAACANACTCNGCAAACAA  
GGAANNAAACGNGTTNNCCNTAAANACATGTTCANNGAATCGAANTNCAATAACTGTNGACTCC  
TGAAGAAAAAGTTTCNAAAATTTATNAAACANGCCCTNGTAATTACTCCNCCANAAANNANCTTG  
GCCTNNCCCNAAATTTATTANAANATTTACCANACCAGTTGGTCCCGAAGATATTTTTTTN

SEQ ID NO: 663 ACATTAAATGTTACTTTGGCATTCTATTTCTTAGGCTTACAGGAATTATTAA  
GAATTCCTTTGTAATGCAAATAATCACTCTTGGAAATTAATTTAAGTAGAAAAATGTTACATTT  
TAAGGCAAGAAAAACATTTGTAAATATTTTTATAAAGGCATTTAATTCAACAAATTAACGGAATCAT  
AGTAAATGGATGCAAATAAAGAATATTACTTTATGAGAAAGACAAAGTTCTTTGGAGCTTACTGT  
GACAGTATGGAAAACTTGTCTTCTTTTTTAAAGATAATTCCATTTTGATACTTAATAAACATAAT  
GTTAACCTTCAAGAAGAAATAAGCACTCTTAGTTGAACAAAAAATGTAGGGAAATTAATAACAA  
AAAATGAAATTAAGAAGAAATACCAAATGTNAATTGCCTCAAAATTTATCTTCTTCTCTANAAAA  
AAGGATATTTTCTCAGGAGTGACTTAATTGAAGCCCTCTTATAATTTACATATGGAAGTGGATGGT  
AGAGCTCATGGTGGCAGGGCATTCCATGATTAAGTGTTAACAGGGNTTAC

SEQ ID NO: 664 ACAGTCCGGCCCGGTGGGGAGGAGGGAGGGAAAGGCAGGCACACGAAGACAC  
AGGTATGTCGGGAAGTGCACACAAACCGTTGTCTTTCTTTTGGTTAAAGAANAAAAACTTTGTA  
ATCAATATCCTGCTCATAAGTAAAAGTGGAAGAAAGAACTTGATTGCTTTCATCTGGCGTTTGTG  
GCATCTCCTCTCCCATTTTCATATGCACAGTTTATTTGGGTAATGCTACCGTCACCAGCAGAACACCT  
GTAAGTAAAAACAAATGTCAGGAAGGAAAAAGTATGAACAACAGGAAGCTCCAGAGGCGGCTCC  
ATGCGGGCGCTGGGCTCANTAGAAGCAAACGGTGTGCAGAAAAGGTGCCGGCCACTCTTCTCCTC  
GAATCCTGCAGCAGCTCGTTCGATCTCGTTCTCCATGTCTTCCGTGTGCTGGATACACTGGCAGAG  
CTCACAGATGAAAAACGCCCCCAGGGTGGCTCCTCGTGGTTGTCTGGATGCCACATTGACACNT  
NCACAGCTGGCAGGTCTCCTGTTCTCTGGAATTAATTTCCACCACCTG

SEQ ID NO: 665 CGTCAAGTTCTTCTAGCAGGGACCTGTCTCCCTTTACTTCTTACCTCCACCTT  
TCCAGGGCTTTCAAAGGAGACAGACCCAGTGTCCTCCCAAGACTGGATCTGTGACTCCACCAGA  
CTCAAAGGACTCCAGTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGT  
CCCTTGATAGGGTGAAGCTGAAGCACCAGGGAGAAAAATATGTGCTTNTCTCGCCCTACCTCCTT  
TCCCATCTGACTGTCTTGANCCANGGTCTGTAAACCTGACACTTTATATGTGTTACACATGT  
AAGNCCTGCCCGGGCGGGCGGTGAAAGGGCG

SEQ ID NO: 666 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTATTCAAATAAAATTTTAATGAT  
TTCAGTTCCTGAAAATGTAGNGNCATTAAAGGNCAATTCCTGCTAAATTTCAAATTACANATTTGN  
GGGCCATTCCTGANCAAAAGCATCATTTCTACTAAATATTCAGCTTGNTAACTAAGGNAACTGACA  
GTNTCATGGCAGGAGGNGANAAGGAGCAATGATCAGCTCATAGCTNAAAAAGGGAAAAAACA

NTAACCAAAGCAAAAAACAAAATAGCTNCNGATATTNGTAAAAATAATAACAATCCCAAACATC  
CCCTNTTAAATTNCTACNAAAACAATACCCATTAAAGGTTTAGCAANTACAGCTTCNAAGATTTGAG  
NAACTTTATTGCATTTTATAGNGATTCTTAAGGCCTATTTCCAATGAAACCATTTTAAAAAGCTCTAT  
GAGGAGNGGAATTTTANATGTCTATTACACTTGGCTTTTAAAAAAAATGCTTAAATTTCAAAAT  
GAGCAAAAATCACTTTTGGTNG

SEQ ID NO: 667 ACCTGATTTTTTTTTTAAATGGACAGTCTATGCTTTCATGAGGCATGCATATTG  
AGCACCCCGTGTGGGATGGTCCCAGCTAGGGGAGAAAGTATGTGTGTCTGTGCTCACCCGGCAGC  
ATAACTACACAACCTGCCCTCTGAAGGATGGAGCAAAGGACATCTCCTGAGCACTGAGCATCAAG  
CACTGAGGCTCTGGGAAAAGGAAGATCAGACAGAGAGGAGGTAAGTGGAAGGAACTGAGCAGCT  
TCGGTTCAATCCCAAGGCAAAAGCCTGAGCAAGGACACACTCCCAAAGTGCTCAGGGTGGTTCCC  
CGGGCCCCAGGACGGGTGACACANGCACAAGCACAANCAANNCCTTGAAAGCTNTGAGTTTNTT  
GCTTGAGTCTAATTATCTTTAAAGGAAANAGACTGGATCCTCAAGTCTTCCAATCCATCCATTGTA  
GAATACAGGACTTCTGTCTGATGTTTCAAAGGATGAATATNTTGAANGCTGCCTCAAAAGACTTCC  
CANTCTGGGCANACANTNNGACCTTATTTCTGTAAAAAATAAAAAAAAAAAAAA

SEQ ID NO: 668 ACATTGGTTGGGGGCCATGGCGGGTTCCACATTCTTATGATTCTAATAGTCAG  
GGCGGGGCATGCTGCTGTCTTGAATAGGGATTTTGCATCTGGCTGTCTCTGGAGGCCGTTATCC  
TGACCTTAGGAAAGACGCTCCCTTCTGCTCCAAGATTGGTGAAGCCTGTCNACCGCAGTTTATAGAT  
TTTGAAAGTCTTCGCAATCAAGTTCTTCTTCAATATGCTTCGNTTGTATTATAATCTTGCTGTG  
ACAATCAATGCAATAATCATGNTNANAATAACNATGCCANCGANNGTGCCACANTAGTNANNAT  
CACATGAANTTTGCCTTACA

SEQ ID NO: 669 ACAAAGGAGCCTAGGAACCCCTCAAAGACCCTTTCCCCTACTTACTTCCCACA  
AAAAAGAGCCTTGACAAACCAAGTTTTATTTCCAGCAACATTCAGCTGATACAAACACATAGAA  
ATACAATCTCACCAAAGAGCTAAGAAGCCCTCCAACACACGCCCAATTTAGCATACAGCTAAGA  
CTACACAACACACACACAAGCTCAACAAAGAGCTAAGACCACTACACACATCTCCACAGCACCAC  
AGGCACATCTCAAAAGGGTCTGTCCAGTAGTTCAGCCAAGGGAAGGGTGATGGCTAAGCCACCA  
GCTCCAGCTTTCTTTCTTTCTTTTGTAGCTGGAGTTTGCTTTGTTGCAAGGCTCCTGACAGTT  
GGACACATGGCTGCACACATGCTGCTACTTGGATCCCAGAGGGTTCTCTGGTGGCAGNTTGACCN  
ACTCTGNCTAGGNATC

SEQ ID NO: 670 ACCCGCCTGCCATGGACTGGATCTTCCAGTGCATCTCCTACCATGCCCCGAG  
GCTCTGCTGACCGAGATGATGGAAAGGTGTAAGAACTAGGAAACAATGCCTTGCTGTTGAATTC  
TGTGATGTCTGCCTTCCGGGCTGAGTTTCATCGCCACAAGGTCTATGGATTTTATTGGCATGATTAA  
AGAGTGTGATGAATCTGGTTTCCCCAAGCATCTTCTTTTTCGATCACTGGGATTAACTTGGCCTTG  
GCTGATCCTCCTGAGAGTGACCGACTTCAGATTCTCAACGAAGCTTGAAAGTTCATCACTAAGCTG  
AAGAACCCACAGGACTACATTAATTGTGCCGAAGTGTGGGTGGAATACACCTGCAAGCATTTCAC  
GAACGAGAGNNNATCCGTTTNGCAGATGTCATCAACACATGACTCNNATCGGCTTTGAANATTCT  
ACCCANTTTCANTATATTAAGATATTGCNCTTCATGCTTTTCAATTTTNTATGAAAAATTTGCGTT  
TGNATGTCCAAAAANANGGNGGNGAGGTTNAATGCTATGGACCCTTNTAAAAAAGAGCC  
AANGCCGNTTTAAGCCTTNTGTTGAAAATCTG

SEQ ID NO: 671 ACGCGGGGACAGGAACAAAAGCAACCAATTTTAACTTTCTCTTCTCATTCT  
GTTTTCAATTGATTTCCACATGTAGTCCTTTTGCTCAGGAAGTCTTTGGGGAATTAAGGATCTTTG  
AAGCTCTGAAATAGGTGATCAGGTTAGTGGTGTCTGTGAGCTGTCTAAGAGGTTGGAAAATGAAC  
TACTCAAGATAGTCACGAAAATACTGAAAGTTTGATTTTCTTTCCATATTTGAATTAATTTTTCT  
GTTTGACTGGAAGGGGTTTTTGTATACTAAAACCTCAGCGCATAAAGGAGATTTAAAAGGAGCA  
CATGATTTAGTGGGTGGNCATGAAACTAGAATGGGATTTGGGGGNGAATTTGCNAATTCTGGATT  
TAATCCAACTCTNTGCTACAAGCCTTGGAANGNCTTNAATACTTTNCTCTTTTACAANAGAGGC  
TGCTTAATTATTTGNAANCCCTTCANGCCTNATTCACAANTCCTGNCGGCGGCCGTNAAAGGGNG  
AATTCACCATGNGGCTGTATTAT

SEQ ID NO: 672 ACTTTTTTTTTTTTTTTTTTTTTTNGGNGANATTATTTACTAAATAATTGATAT  
ACATCCAACATCACTGAATGGAAATAATTTAAAAATAANCAAGGCTTAAATNGGGGTCTTTCAA  
AACTCAAAACATTCTNGGNGACTGTNTCTTTAACAAANGGAATGACAAAACCTAGTTAAGTCATA  
TTATCTCAATCAACTCCAAACTCAATGTTGCCAACATGTTTTNTACTTAAAGCAAACAAAAAG  
CAAAGCCTCATAATTTAGTAATTACCATGTTTNTTGGTTTCTGTNCTCAAATGCCAGAAAGATCC  
TGTGAAAGGAAGCAAAGCACCTANTNTCATCAGTCTTNTAAAACTTTAAGNTGCCTCCAACTCCT  
TAAATTTANATCCTNCTGACAATTTCTCCTATGGGAAGANGAATGCCTCGAACCCTAGTTTTCT  
TTGCTTATCNAAGCTGTTNNTAANCNCTTGNTTNTGCTGCACACTAAAAAACTATTCTATCTTTC

AAAACCAAAAAATNTAATTGAATTT

SEQ ID NO: 673 ACACCATTACTCTTCTGAAGAGCCTGGTAAACAGCAGATGCTTTATTTCAGAA  
AACTAGAAAGAGCAAATACTCCTGCCACATCTTGATGATTTCTATATGCTAAATGCATTGCCATGC  
ATCCTCCCATAGAGAATCCTCCTATTAATATCCTGTTCTTCTTGATGCCACTTTTTACTTCTTCATCA  
ATCAAATCAGTAAGCACTTGACACATGACATCAATTGATTCAAGGTGTTCTGGGCAGTCATTGGTT  
ATTTTAAATCTGTCAAACCATACATTGGAGATTCTCCTTTTCATAGGAGTATATGATCTGGGAGGA  
GCTGTTGGANAATAATTTTATGTGTGGAATGTTAAATCTTGATTAAACCTGNTGATCCACATCTT  
AATCCTGGCCAGATCACTGACCATGCAGGAAGACAAAGAGGCCTTTGCTCCTGCGGAAAACNTCA  
GCGTGANAACCNACCCGACGCACCGCATGNTGATCCCCGGTNGTCCGCGCGTTGAANGCNAAT  
CNNATGCGGCGTATNTGATCGACTGGACANCTNGGGATTGGGTANNTTCCTGGGAATGTTCCNT  
NNANTCCACACANACCGGAGATAAG

SEQ ID NO: 674 ACCTCAACATAACCTGTAAAAGTATTTCTAGATAAAACTTTACAAGTGAAGA  
AAGAAAACCCATGATGTTACACTTACACACACACACACACACACACACACACATCATTCTTA  
AGGAAGAACAAAAACATGGTAAGAGTGTGAACAGGAAGGGAATGCATCTTTTTTTGTAAAGCTTA  
TATTAAAAAACACAGCATGAGTAAATAAATTCTNTATGCCAAGAGAAAGATGCAGAGAGAGGAAA  
CAGAANGCNGGGATGAGACTAACTCATTAATAAATAGTTNGAGAATGTCATTCAAAAAACAGTAAA  
TTTGGGGAGTTACANATAATCCCCAGTTGCACCACTTNAATTTTTTGTTCGATAACTATTGACCC  
TAAAAAANNCTTNTATGNGTNNTTTTNGGGANGGNTTNGATTGNGCCNNAATTTATTAAGGGGG  
CCTNNGCCNTTTNAAGGGNTNAATCCCCAANTCACNCNNANGGGTTTTNNNNGGAACCCCCAA  
AAGGCCCCNTTTTTAGCCNAATTNTNACCCNNGTTTATNTTGGGGGGGCNCTNNGGGTTTTNNA  
GGGNAATAANGNGGCCATATTTNCNTGNTTGGNNTTTAAAAATNATTGGNTGGNAAGTTTC  
NGGCCCC

SEQ ID NO: 675 ACCAGGCTGGCGACAGGTGCTACCAGGAGTGGGCTGAGGGGAGAAAACTA  
TCTCCCACTCTTTTGGCCAGGCAATGTCAACGACTTCCACATTCCCTGGCCCACTTCTGAGCAA  
CCCCAGGTTCCGCTCTGTATAAGGACCTCCCTCCCAACCCCAACCCAGAGTGCAGTGCATAATC  
AACCAACAATTTACTGGTGAATGGCAATCAAAGGAACAGTTAAACACCAACAATTTCTTAAA  
GCCAAAAAATATTTTTTCATGGAGTTGAACATTTTTTCGAGTGTGTTTTTTCAAGTGTAACAAACAGT  
GACATTTTGTTCAAACAGAAACAGCATCTANGAATTCTGGCACTTGGGTTCTAGGGGGTTACAGGT  
NTCATCATGGATTCTTCTCCTTGGAATTAATAAGGCCTCGNGTTTCTATTCTGANTTATACCAACA  
CCTGCTAGCTTTCCNTTTTAGCGGACAGTGGGTGGGCAACCAGCCTTCTGGTTANATGGGCAATG  
CCAANCAAAAATTCCTTATTCACTTGTNGGCTTGGTTTTTTATTCAAAGNAANTCAAATNCCAAA  
AANAACNTTACAAACCCAACCCCCAAAGGCGCCTTGACNGGGACCCTTCAAACCTGATTGGTGG  
AA

SEQ ID NO: 676 ACAAATTTAAGACTAGACAATTAACAAAAAACAACAAAATTATAGCCGC  
AACACTATCCTAGCTGCATTGTGGTGATTATATTGTTCCGTTAGCTGTCCATTATATACAACCTGGAT  
TTTAATCGTGCTATTCACAGAGGGGCCAAAGCACTCATCTAAAGGTAAAGGCGCTCAGGTTAAC  
GTGCACTTTTCATAGCCCCATGAAAACACAACCATTATATACACTGTAAACACCACGGGGCATGAA  
ATCCACCTCCAGCCACAAACATCAATCTCACAGGATCAAGCATAACGAGGTCTAAGGTGGGGTGCC  
TGACTTTTCTTTTAAAGAATACTTTTCATCTAAAAGTCTGCTATTCTTTTCACAGGACAATATTCA  
CTTATTTCAAATATGTTGCCAGTGCTCAAAACATTTCTGGAACCCATTTTGGAGACTACCTGCAG  
CATCTGCAGTGAACCTCTTCTGAATGTCCTGGATGGTGGTGAAGTCTTCATCCTTTGAGGATGGGTTT  
GATTTANGAAAACAGGCAGTTATTCACAGCCAACCTGAATCCAACCTGNGGGGATCCCCGGGTTTC  
AAACATTGANATGCCCTATAAACACAAGACCAATTTTTGTACCTCGGNCGNACCACGCTAAGGC  
G

SEQ ID NO: 677 ACGCGGGGGAGGCCGNTNTCTCATCGAAGATGGCGGCGCGATCTGTGTGCG  
GGCATTACCAGAAGAGTCTTCATGTGGACAGTCTCAGGGACACCATGTAGAGAATTTTGGTCTCG  
ATTCANAAAAGAGAAAGAGCCAGTGGTTGTTGAGACAGTAGAAGAGAAAAAGGAACCTATCTTA  
GTGTGTCCACCTTTACGAAGCCGAGCATAACACCACCTGAAGATCTCCANAGTCGTTTGGAACT  
TACGTTAAAGAAGTTTTTGGTTCATCTCTCCTAGTAATTTGGCAAGACATCTCCCTGGAAAAATAG  
TCCGTCTAAAGTTCAATCTTCTGCTCATTTAACTGATGACTTGGGTGATGNAGTCCCTAACTCCANA  
CTCCACCANATGTGCAGGGNTAGAAATGTTCTTGATTTCTATAATGGCCCTATTCAAAGATAGATC  
TAAATTTGATGAACTCAGTGCCANTAATCTGCCCCCAATTTGAAAATCACTTGGGAGTTCTAACA  
ATTTTCGGAAGAAAAACCCATTTGAAAATCACTGTCTTTCCCTGAGCAANGGGGGCTGCTCATTA  
ATCTTTTGATACTTTACCATGGGAAATTCNCCANAACGTGTTCTTTAAACCCACTTTTTNTGTAAA  
AG

SEQ ID NO: 678 ACTTTNTTTTTTTTTTTTTTTGGCTTTATAAGAGAATTTTTATTGTTAATTATT  
TACCTTAATAGTTTCAGAAAGAGGAACAAATTACCTCAGTCCAAATTGATTGGCAGTTGGCTTTTT  
CTAGGGAACCANNGGTTCTGACTGCTAAGGATTNATTTGGGATATTTTNAATACTAANCCTTTTAA  
CCTTNAGGCTTACCCCANAAATAATTGCTTCACCTTCCTTTTCANTTAATTACCGCNCCTACCTACT  
TNGTTNTANAANTNTCNTTCATTTTGATTCCAAAAGGCCAAGTTTCTTGAAAATACAGGCTTGAGG  
ACATGGCCAAAAAANAAAACCTCAAAATAAACCTCTGGATGATAATTACTAGTCTAAGGAACCAA  
ACCNTNCATTNTNTTAANAATAATTCCNTTNCAAAATGCCCTAATAGGCTNTGTGAAAAGGTTTTG  
GTNTAATAACTGAAATGGGTGGCTTTTTACAANATCCNAAGCAGNTTCCGGGGCCCCCTNANCCAA  
CANTGGCCCTNAAAACCAAAAGGGGTTTTTCCGTGGNGGCCATTTTAAAAATCCNCTTCATCTT  
TTCTCTCNGAAAAGCTTGCCCATANTTAGGCTGGACTTGNGGGACTGTNTTATCCCAAACAG

SEQ ID NO: 679 ACCTGGGTGTATCCTGTGTTTGCCAACTCAGCCTCTTGGGTCTAGCAGCTTT  
CTTCTCTCAGCTACGTCTTCATCGCCAGCATCTACCTACTTGAGAGAAGCTCAACCACTGGAA  
ATGGGGTGACATGAGGCAGCCACGGAAGAAGAGGAAGTAATTGCACACCATTTTCCAAGAACCA  
AGAAAGAAGAAAACACAAGAGATTTTCTCATCTTTTTTTTTTTTTTCTGGTGGAGGGAGGTGG  
TGGAGGAACTTANCAAAGTAGGAGGGACAGANAGTGATCCTTAAATTTAATAANAGTTCGTGAAG  
GTAGCTTAACCTGANAACTNTTGGTTTTTTGAAAGGTTGCTGN

SEQ ID NO: 680 ACAAGATATTTGCGGTTTTGTTTTTTATAACCCACTAAGCCAAGATTTGTATC  
TCTGTATGGAAGCTTTTTCAAATGGACAGAAATGGTCTTTTGATCTTTCTGAACCACTTGTCTTCAA  
ATTCTTCTGAGGATACAGTCACCAAGGCAGTCAGGGCTACGGANCCAACACACTTACCTCTGGGT  
GAACTCATCTTTTATTTTTCTGGGATATCTTCCCATAACTCAGCTATCAACAGCAAAGTGTCT  
TCTTTGAAGCTGAACCTTCATTTTATCTCAATTCTGCTGAGTCCTTAAGTTCCTCTCCAAACAAA  
CGATGTTTGCTGAGGTTGGGTAAACTGAAGCAGAACCTGGTTCGTATAAAAAANGCCATGCTTTC  
AACTGAGTCTCTATCAACACTGTTCTCTCTTTCTGNCATACAAGTTTTGGATATTTGNCANAGCTT  
TCTCATTTTCTGGCTATTACTGCTATTCGCTTCACANAAAATAAAAGCTCTCAGCAAGGTTCCANG  
AAAACCAAAANACAGCTTTTCTGAGTGTGGAAAAAANGATTTGATAGTTTTGTCTTGAGGTGA  
ACTGNGCTGTTTTGAAATGACCCAAGTGCTGCTGAAANAACCTGGCCCCAATTCANGTAANCCT

SEQ ID NO: 681 ACAAATATCATCATTTATTTTTGATTTTTTTCACCAGCCCTGAATTTTCAAA  
CTTGTAATATGCTGTTTCACAATCTTTTTATTAAATTACTTAATGATTCCAGTCTGCAAAATGAGCAT  
AGACTTTGCTCTGTTGTTATAGATCATTGAATTGGGGGGGAAACATAGTAAGCTAAACTATCCGTC  
ACTGCTCATGAAAGACCCACATTGTTGATTATTTTTCCCAGCACAGAAAGAAAAGGACAAGTGAT  
TCATCCTGCAATGGGCCTCTGCCAATCTTTTCATATCTATGTAACCTTTTGATGTTACCTGGTGTGA  
TTAACCGTTGTGCTATTGTAATCTTGGATTAATTAACAAAACAACCAAAAATCTATCACCTGGAA  
TATTGAAAAGAAATTCAGTAAACAAGATGTGTCTCATAGTTAAGGAGAGACATAAAAAANTAAAA  
ATGTCATTNAACAGGTTGGATTTAGGATTTACTGTCTAATCNAGAAACACCGAGGAGGCTTANCTCA  
CCCTTTNATTGGAGAATGTGGGGAANGGAAAAGAGAGTAAACACATTAACTTTAGTTANCNANAA  
GTGCTGCNTAAAAAATTCGCTGAAAGGAAATGGTTACAAGACAAATTGGCTTTTATCCCCCTTGNT  
T

SEQ ID NO: 682 ACACATGCACATCAAAACACTTCAACTGAATATAGATGCCATTACATTATTTA  
GTTACGTTACAAAGCAAACGGCAGGTTCTATAACGTTGTTCTATTATGTATCAACTGAAAAAATA  
TATTCAAAAAAAGTTTTTTGAAAGACTCATGGGGAAGTGGAATGGTGCCCCACATTAGGAAAT  
AAAGCTTTTACCAGGACCACCTTGCTCCAGCTGGCTCCCAAGGACCACCTGAAAACAGCTGGCTT  
CCCTNAAAAANACAGAATGGGTCTTGTTAATAATTCAATGGACTCTCAAACCTATTCTCTGGAT  
CACAACGGGGGNCCCCAACCTGATAATGGCCCCATGGGTTGGCTTGCCAGAAAGTCATTATCTCG  
AANTCGNAAACCCCTTTCAAGCATCCAATCTTTGGTTTCTTTAATGACAATAGCGTTAATAATCC  
TTTNCNTNTACNGNANTTACAACATCANAAANGNACTTCATGAGTTCATTTTTCAAGAAAATGCCA  
ATTTGGCTGNATTTTCAACNANGGTTTTCTTCTTCTAAAACCTNAAGGGCTGNNAANGGCACTCGG  
NACCCTATGGATTGGCACCGTTTGTGGCCCATGCTCCCTGCTTAATGGTCANNATGATGCATTATN  
C

SEQ ID NO: 683 ACTTTTTCTTTTTCTTCCTTTTTTTTGAAATTATTTTCTGAGCCTTTTGT  
ACGGTATATTGTAAACTTTTATGTTAAAGAAAAAATATACATTTACAAATTGTGAGATTTTAAAG  
GAAATTTTCTACGATGTATACTGGCTTATTTTTTAATTTAAAACGGGGTTTCCGTGCGCACTGGTGG  
AGGGGGTGGCTGTTAGTCCCCTCGCTCCTGGCTTTGGGGGTGGACTTNGGGGGNCCAGAACTC  
TGGNAGCTTCTANAAGAAATCTACTGAGTGGNNTTCTGTTTTTGGTTAAATCCCTGGCTTTGNCG  
ACTGGACCTGCTTGAATGTCTGAGGGNAACGTGTTGGGGGTGCNACAGCCAGCCCGGTGGATCCAC  
GCAGGCTTAACCGACCNATAAGAAGCCTTCTCCAAGCACCGTGGTTCAGGGCGTTTCCATTGACC  
AGTTTGACCCTGGTTGAATAAANAATAATGCGTTNGGTTTNA

SEQ ID NO: 684 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAATTGAACACATATTTAATAATTG  
TGATAATGGTTTTGNGGCTAAGTTAAAAAATGAAAGGCCCAATCTTCAAAAATCTACTCTGAAA  
TATTTACAATTACATATTTAACAAGTNTATAAAAAGTATAACCATATNTGGGATTTGCTTTCACTTA  
ATAANAGATGGTACCTGGANAGATGTATGGATGANATAANATTGGCCACAAGCTGATGGTTACTA  
GAACTAGGGGCTCGAAAGTTAAAATTNTACCCAAAAAATNTCTAAAAAATTTTNGNCCANAATT  
TTNAATCCCATTAATTCAGGAAAAANANTTTCAATNGCTTAATANCATTTNCAAATAATAAAAAAGT  
TTTTGGAAAAANAACCTTTGGNTANANATTTTNTCTGAATTCANACTNTAAAGATTATGCAATNATG  
GAAAAANTATTTTGATTAANAANTAACAGTTTAGGAAACCCTTGATGNNTAAAGTAATGGAACCG  
TTTCCTTACAGGAGTATAAATANGCATAATCT

SEQ ID NO: 685 ACTTTTGGAGGCCAAGGCAGGTGGATCACTTGAGCTCAGGAGTTTGAGACCA  
NCCTGGCCAAACATGGTGAAATCCCATCTCTACTAAACATTNAAAAATTAGCTGGGCATGGTGGTG  
CNCACCTGTATTCCCANCTACTCGGAANCTGANGCATGANANTCGTNTGAACCTGGGAGGNNGA  
NGCTGCTATTACCTAAACATTGTCNANTGNATATNCCCTATGCANAGA

SEQ ID NO: 686 ACGCGGGGGCTTCCGGTTCTGACGGACGCTTCGGCCGTAACGATGATCGGAG  
ACATCCTGCTGTTCCGGGACGTTGCTGATGAATGCCGGGGCGGTGCTGAACCTTTAAGCTGAAAAAG  
AAGGACACNCAGGGCTTTGGGGAGGAGTCCAGGGAGCCCAGCACAGGTGACAACATCCGGGAAT  
TCTTGCTGAGCCTCAGATACTTTTGAATCTTCATCGCCCTGTGGAACATCTTCATGATGTTCTGCAT  
GATTGTGCTGANAAAACTCAAGCATATTGCCTTTCCATCTAGCACTGGGGCCATANTTCTGATAC  
ACTGGNAACGAAATTGNGAGATTTGCTGTAAATGGATNTAGGAAAAANCCNAACGGGANATTTGT  
GATNTGGCCTCTNGGGTNCATCCTNGTNNCTACCCAAAAAAGAGGCTAAAAANTNNCTNTCNN  
NTTAATGATATTCCNNTGTGGTNGGCTGCNATTTGTTNCCANCCAGTNNTTTTGTAAANTNTT

SEQ ID NO: 687 ACGCGGGGGGGCGCACCCGCCGATTGTGGCCATGGCGGCCGAGTCTCTAG  
TGTGGTGAGACGAGTGGAAGAGCTCGGGGATCTGGCTCAGGCCACATACAGCAACTTAGCGAAG  
CTGCCGGTGAAGATGATCACTTTTTAATTCGGGCCTCTGCAGCATTAGAAAAATTGAAACTCCTGT  
GTGGAGAAGAGAAAGAATGTTCAAATCCATCAAATCTTCTAGAACTTTACACACAGGCTTTTTGG  
ACATGACATATTTTGAGGAGAACAAGCTAGTANATGAAAATTTCTGAAACTCTTTTACAGAAGT  
AAAAGACTTGATNANTTTTTTCAAACCAGAAATTTTANTAAAGGAATAATTTCTNTCCAAACAC  
TGCATTTTGTCTGGGGATGANCTCTGGAATGCTTCTTTTGGAACNANGACCCTGCTGNTATGTATT  
GCATTCTTTACCAAAAAANANAGTGGCTTTAAAAAATTCATTTGCTTAAAN

SEQ ID NO: 688 ACGCGGGGGAGACCTGGCTGCTGTGTCCCGCGGCTTGCGCTCCGTAGTGGAC  
TCCGCGGGCCTTCGGCAGATGCAGGCCTGGGGTAGTCTCCTTTCTGGACTGAGAAGAGAAGAATG  
GATAAGCCCTCTTCCCATTAGTGCCCTTGCATTGGTTTGCTTTGNNTCACANCACTGG

SEQ ID NO: 689 ACGCGGGGTGAAGATATTATGGCTGCTGCCACGGAGCATAATCGCCCGAGCA  
GCGGTGACAGGAACCTGGAGCGAAGATGCAGCCCCAACCTCTCCCGAGAGGTGCTCTACGAAATC  
TTTCGCTCCCTACACACCCTGGTTGGACAGCTTGACCTCAGAGATGATGTGGTGAAAATTACAATC  
GATTGGAACAAGCTCCAGAGCCTCTCGGCATTCCAGCCTGCNTTGCTCTTTAGTGCATTGAACAA  
CACATTTTATATTACAGCCTTTTTTANCAAACTTCAGTCTCCGATTAAAGAGGAGANTACAAC  
GCTGTTGAANAGATAGGAAGAACAAGATTGGGGAAACAAAAATGAAGTAAATGACAAATTTTCCA  
TTGGGCGACCTACAAGAGGAAGAAAAGCNCAAGAAGGTGATTAAANAGATGTAAAAAAGACAC

SEQ ID NO: 690 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCAAAAACCGCCAGTNTTTTATT  
CTCATGAAAAAACANAACAAACCCACAANTTGGAGTCACGGANATAAAATACAGATGAAATGGA  
AAACGGTCTGTTGTCAATNAACNTCACTTTCAAATACCATTTTATATGGAAGTTACTTTACTGCGG  
GGCAAACANAAGGCCATGCTGGAGTCTNTTACTTTTGAAAAATGAAAAATCAAAAAATTTGCTANT  
CAACAAACAAAAAAGGAGGGAACCTCCTTGGTAAAGCTTTACAAACATAATNTCATTNTNTNTT  
ACCAATAAAAAATAGCTAGGGTAAAAAAAACAGATGGTTAAAAACTGGTGCCAAACCAAAGNGA  
AAGCTTTGGNGCCTTCNNTAAACNCCTATCCTNTTTNTTTAAAAAACNCCATTTTTTTAAAAAAG  
ATTACTNTGAATTCNAGGGTTTTNTCACCCCAATTNATCCCTNCCTTTTACCCCCAANANACAAA  
AGGGCCATGTTNTTTTTTGTCCGGCCCTTGGGAAGGGCTTGGCCTNAAGGGCTGGGAAAAANCTC  
CTGTTNCCCTAANTTGAAANGGANGAGGGGGCCACCTTTTTTTTTTTTTTTCNCCCTTAAAAA  
AAAAA

SEQ ID NO: 691 ACGCGGGGGGGATGCGCTTGGGCTCCCTGTTTCGTTCCACATGCAGGGCAGC  
ACGAGGAGAATGGGCGTCATGACTGATGTCCACCGGCGCTTCCTCCAGTTGCTGATGACCCATGG  
CGTGCTAGAGGAATGGGACGTGAAGCGCTTGACAGCGCACTGCTACAAGGTCCATGACCGCAATG  
CCACCGTAGATAAGTTGGAGGACTTCATCAACAACATTAAACAGTGTCTTGGAGTCCTTGATATTG  
AAATAAANAAGGAGTCACGGAANATNATGGGAGACCCATTTATGCGTTGGTGAATCTTGCTACA

ACTTNAATTTCCAAAATGGCTACGGATTTTGCAAAAAATGAACTGGATTTGTTTANAAAGGCTCTG  
GAACTGATTATTNACNCAAAAACCGGCTTTGCNTTTTTCNAAAAANATATTGAACCTTGGTTNATCA  
ATTAAAGGCAA

SEQ ID NO: 692 ACCCCACACCCTGAAGGTGTCTATGAGTTCACATGGCTCAGGAATGGAGTTTT  
GGGGCCCTAGATGTAGACAGCTGTGGCGATGGAGGCCAAGGTTTTGGCCTGCTGTGGGGGTTGCA  
TTCACCTTCCCATCTCCACATGACGATACCCAGCATTTACGAGGCANCTTCTGCTGACCTGGCCA  
AGCACCGTGGTCCNACTTCAACGCATCTGAANCGTTGCCAGGGANCCCTTTAAAAATGCAAAT  
CCTTGCCCATCCTTGAAAACCTTTAATTTTGTAGGTNTGGGNTAAGGGGGGGCCTTNNAGGTTTTGC  
CTAAAAACAATNTAATTTGGCCCCNNGTTTAAAGNATAACATTTCTNAAGCAC

SEQ ID NO: 693 ACATGGTAAAAAGTTGTGGCAAAGATGGCTTCCATATCCGGGTGCGGCTCCA  
CCCCTTCCACGTCAATCCGCATCAACAAGATGTTGTCTGTGCTGGGGCTGACAGGCATGCGAGGTG  
CCTTTGGAAAGCCCCAAGGCACTGNGGCCAAGGTTACATTGGCCAAGTTATTATGTCCATCCGCA  
CCAANCTGCAAAACAAGGAGCATGTTATTGANGCCCTTCGCAAGGCCAANTTTAAATTTCTGGG  
CCNCCAAAAATCCCATTTAAAAAANNGGNTTTACCANNTTNAAAGCTT

SEQ ID NO: 694 ACGCGGGTAGGCAGAGAACAAAAATGTTAAGCATGGTGTGTCTATCTTATT  
GAAGCGGTTGGAAATGAAAGCTTTTAATTTGATAGATTTATCAGTATAAAATTAAGGGAAACCAC  
GTGNGGGGAATGAATCAATTTAAACTTCGGGAATTGNGANGNGACTTTTGNAACTTTTGGTCTG  
NGTGTGACCTGTGAACCACTAGGATGTGATCTGCCCTTGNNGGCAGGTCCAGCATANTTAGGAGT  
TAGGCTTTAACATAAATTTCTAGCTGCATCTGANTCTCCTGGGATGGGTGCTCTTTGGCTGGTTTTG  
GCTGCGGATGGTGANATCAANCAANTNTTCTGNTGNTGGCCCTGCAATCAGTTGTTGGGATGC  
CANTGCNAATCACTTAGTNGTNAAATTTTAATCAAACACCACCAGGTCCCAAATGCAGGTCATGA  
ATNINAAATTTCTAAATTTACATNAAAAGTNAAGTTTNNACCGNTTNACATTTTGGTNTTTAAGGA  
AAAGAAATTG

SEQ ID NO: 695 ACTAAAGGCTTTTGCATGAATTAGGAAGGAGAGTCTTGGGGCAGAAGCAATA  
GGGGACAACCTGTGCTGGTGTCTTTTGCAGGATGTGTTTACCAAAACATCTAATGCAACTATTT  
TCAGACTTTACAGTTTGTAGTGTTAACTCTTTAGAAAAAGAGCAGCCATCCTTTTTAGGAANAC  
TGTATCATCCCCAGTGATGAGGAGAAGCTCTTCTGTANANAAGAATGACACTGTGCTGGGGCAAG  
CGATTGCATACTGNANCGGNCNCANGTCCCCGNTACCTTGCCTTNTANNGGACAAGGGGGCCCTG  
GACCTTTCCCANCCCTTTG

SEQ ID NO: 696 ACGCGGGTATTGAAGAAGATTTTAATATTGCACTAGGAGTTTTTGTCTTAGCT  
GGACTAACAGATTTGTTGGATGGATTTATTGCTCGAAACTGGGCCAATCAAAGATCAGCTTTGGGG  
AAGTGCTCTTGATCCACTTGCTGATAAAATACTTATCAGTATCTTATATGTTAGCTTGACCTATGCA  
GATCTTATTCCAGTCTACTTACTTACATGATCATTCGAGAGATGGTAATGTTGATTGCTGCTGTTTT  
TATGTCAGATNCCGAACCTCTTINCAACACCNCGAACACTTGCCAAGTATTTTAAATCCTTGCTATG  
CCACTGCTAGGGTAAAAACCAACATTTCATCGCAAGGGGAATCCAGCCAGNCCANTTATCTNNG  
GGCAGCTTCTTTGGCAGCTCCNGTTTCAACTATCCTGANNGNCTTTATCTTCNNATCCTNTGGGG  
NTTTACAAGTTTACCCCAGCTGGNTCACCTTTTAGTTCCTATNTTANTGCCGGGAAGACTGGTTC  
AGGNGATAAAAANCCTGATNAAAGCCATCCCTTNTGTNNGNAAGGGACCNNCNTTCTTTNATT  
GGGGACNNGGGGCCCATTGGAATGTCCCTGCCCGGGGCGGGCCNGTT

SEQ ID NO: 697 ACTACAATGTTCTATGCATTTCTTCATCCTAGACATTAATAAAACACATCCCT  
TTGGTCTTAGATACTTCTCTTTGGTCTGTGTTTTCTTCTTCTGAATTTAATCTTCTGNGATGTGAG  
GAAATTTACGTGAACCTTTACATATCTATTTTTCTTGGGCCAGTTGATAATTTCCCTCCCTTAN  
AATNCCAGCGTTTACCAACTCCTTCGAATTTCTTTGATCTTACTATCCTCTGAGAGGAGAAAGTAT  
TTGAATAGAGCTGGCATCAAAATACAAATGCCACACAATGAATTCTGTCTTCAGCCGATGGGGA  
ATCAATGNAGNCATTGATGAATTTAATTTGATTGATTCCCATGGGATTAACCTGGTATCTATCAC  
CGAATGTTGCCGGTCAAGATATAGAATATGTATCCCTGCACAGGCCGCCTTCTTTNTCACTCAGC  
CCCAGTGAGGCACACAGAATAAACGGNAGGGATTTGGCCATCTTCCCGNCTCTAATAGAAGATG  
NCCTATACTTCTTANATTCCCAGNTGGATTAAAGGCCCAACAAAGCCTGATGCGTACTGCCCTTGAA  
AACANCCTCATGAGTTGAAAAGCTGGCTTCCCACTCCAATTGGACCAGCANCANAATTNGTTTTGG  
TGG

SEQ ID NO: 698 ACTAGGGATACAAAGACTTGGTTATTCTTGTGGAGTAATGATTCTCCTCTAT  
CTGGGGAGGTCATCCTCTGCAGTTGGGGACTGAAGGACTGGGAGATGGGAGATTGAAACTTAAAG  
AGGTAAAGGTATGTCTTGGAAATAAACTAGTTATCAGTCACCACCAGACCATCATGGCGGCTAGA  
ATTCCAGAAACGGTGAATAGAAGTTCATCTTCTATTTTGTCTTCATGGAATATATCACTCTGGTAA  
ACAGCTGGTCTAAATATGATGGGTAGACTTTCAAGGTCTTCATTTTTANATGTAATTCTTNCAAGG

AGGGTGTCTAGAAACCGTGTAACACCATACCTTCCTGAAAATGTTTCTGAATGCCAAAATTGGGG  
GCANTA

SEQ ID NO: 699 GTCGCGGCGAGGTACACAGTCCAGGCTCTCCAAACGGAACCTACAGGCAGC  
CGCCCCGTAAAGTGTACCGTATCGGTTTGGTTCCGTGGGCCAGGACACACAGCTCTGTTTATGGGAC  
CTTACAGAAGATATCCTTTTCCCTCACCAACCCCTCTCAAGAGCAAGGACACACAAAATGTCATG  
AATGCCACGAGCCTCCTGCTGGAAGCAATGGGAACAGTGTTACAACACCCGGAACCTCTGTGCCG  
CTCCTCTGCCACGGTCCAACAGCCTTCCACATTGAGCAGTCTCAAATGCTGGCAGCAAAAAGCAGTG  
TCATGGACGGGGCCATTGCTTCTGGGGTCAGCAAATTTGCAACACTTTCACTACATGACCGGAAGG  
AGAGGCACCACGAGAAAAGATCACAAGCGAAAATCATAGCATGGGACACATTTCTAGCAAGAGCAG  
TGACAAACTGAATCTAGTTACCAAAAACCAAAACGGACCCTGCTAAAACTCTGGGAACGCCCTGT  
GTCTCGAATGGAAGATGTTCCCTNGTTAGACCCGCTGATATTGTAAAAAGATANCNCATGAGA  
GACTTGACTTGT

SEQ ID NO: 700 ACACCGATTGGGATACTTTTCTTTCAGAGCCGGGCCACTGAACAATGAAGTA  
ATCTGAGAGATGAAACAGAATCTTTCCGGACATGGATAACGTTTCTACACGGCAGATGCTTTCAAC  
GTAGACAATGATCACTTTTCTTTATTCCTAGTATCCCAAGGAGAAGGGCAGATACACAGATAGGAA  
CACATGTTCTGGTCCGTTACACAACACCAAAATCTGGCTTCACCCCTGTGAATTAGGGGAAAGGAG  
AGCCACATGGAGTGCAAGGTGGTGAACCGGTGGAGGGCCAGGACTGCTGACCTCCCGGCTTCTT  
GGAATTCGGTGAATGTAGTATTTGGTATACATGTTACTAGGGTCTCTTCAGCTCGATCTANTTCAA  
AAGAATTTATTTATTGGCACTCANTTTCATNATGTCAGCAATGACATAATGTCTAGGTGAGTNGN  
NTTGGACAAAGCTCCAAGCAGCCTNAGGNTCTNNGTGGTTTGGCCCCGGGNNCCACCNTNCCAA  
AATTCTGNAAGACNNCCGGGGCGNNACGTCCNTGGAACGAACCACTNCCCATTTCGGAG

SEQ ID NO: 701 GTACGCCTGTAATCCCAGTGACTTGGGAGGCTGAGGCAGGAGAATCGCTGA  
ACCCGGGAGGCGGAGGTTGCAGTGAGCTAAGATCGCGCCATTGCACTCCAGCCCANCCACAGGGC  
AAGACTCCGTCTCAAAAAACAAAACAAAACAAAATAGATTTGTTTCCCCCTGCACAATCTGGAAT  
TACCATTTCTTCAGAGCACATACAGGCATTTTCATCTTTCAACCTGTAATTTCTCCTAACTCATGTTT  
TCATGTAAATTAACAAACATGCACCTAACAAATAANTNTACAGCATATATCAAGACACAATAT  
AATTTGAAAGTGTGCGAAANAAAATCTTATACAAAATTTAAATAAAAGAAATCATTTTGTCCAAC  
TANTCATCTTATGNTCTAAACACATTATTTGCAAAATATTTTCAAAAGATTGTATGATTTATGTGC  
AAAGTTCTTTAGTGGGCTTCATCCACTGNTTTGGNAGTTTCAAGCATTTTNTCTTACAACAAAGTN  
ATTTTTTTTGTCAATNTTGAACCTATTTCTGATTGCAGNNACTNACTANGTCATAAAAAANATNCTT  
TNTGAAAAAANTNACTACATTTACCCTTAACCTGTTANATTTANTTTNTTAATTNAAAAAACAG  
NTTNGAACC

SEQ ID NO: 702 ACCTGACAAATTATTGGATTCCAGCACAGTGACTCACTTATTCAAGATAACTG  
AAAACATTGGTTGTGTGATGACCGGAATGACAGCTGACAGCAGATCCCAGGTACGATCTNACTCA  
CTGCAACCTGTGCCTCCAGGTTTCATNCGATCTCCTGTCTCAACCTCCCAAATAGCTGGGATTACA  
GGCACCTGCCATCATGCCCAAGCTAATTTTGTGTTTGTAGTAGAAATGGGGTTTACCATGTTAACC  
NGGCTGGTTTGAANTCCCGACCTNAAGTGATCCAACCTGCTTTGGCCTTCCAAATTTGGTGGNAATTC  
AGGGCGNNANCCACCGAGNCTGGCCTTTTTTTNTTTTTTTTAANNCCAATCCCNNGANTNAAACCT  
TTTTNTTANGNCCAATTTNAGATTTTTTTCA

SEQ ID NO: 703 ACATGACCTTTAGTGAAGATTATTTGTCTATCAAATTACCCATATCCAAGTTTC  
CATGGGCCTGGAATTTCTTTCCACTTGATAGAAGTATATATTAGGAAGTCCAGTTAATAATATTT  
TTATTTAAAAAAGGAAAAAAGAATCAGCAGAGTCAAGTGTCTTAGTCTTAAGGCT  
TTCTGGATTTCTTCTTGGAGGAGTTCAGGATCTTCCCAAGGCCTGGGTCCTCGAATATTTCTCCAG  
TCATCAAACCTTGGAGTCTTTGATTTTCTCATATTCGACTCTAAAGATATTTTATTCTTTTCAAGTTT  
TTTTCAAGCTCAGGATCCATTTTACTCTTCACAGCATCNTATCAGATTTGAGAAAACCTCACGAAN  
GACCAAAAGAACCTCCAACNATCAGCAACAACATGGGGACTCCATAGCCNAGAGTNTTGTCTTT  
GCGAAAAGCNCGCATCACCGCGGGTGCAACATTGATTGGAACNTTCCATCNGGCTCANANCTT  
CAAGNGGGCCTTANCGCAGACCTCCNAAATTTAAGAGATTGACGCTCTCNCAACCCNCGTTCCTG  
CCCGGGNGGCCGTTTAAAGGGCNAATTC

SEQ ID NO: 704 ACAGACTGAACAGATTAGGTCTTTGTCTGAAGCTATGTCAGTGGAATAATT  
GCTGCAATCAAAGCCAAAATTATGGCTAAGAAAAGATCTACTATCAAGACTGATCTAGATGATGA  
CATAACTGCCCTTAAACAGAGGAGTTTTGTGGATGCTGAGGTAGATGTGACCCGAGATATTGTCA  
GCAGAGAGAGAGTATGGAGGACACGAACAACATCTTACAAAGCACAGGAAAGAATTTTCCAA  
GAACATTTTGTCAATTTCAATCTGTAAAAGCCAGAGAAGAAGGGCGTGCACCTGAACAGCNAC  
CTGCCCCAAATGCNNCACCTGTGGATCCCCTTTGCGCACCAACAGCCTATCCCAGCTGCCTATA  
ACAGATACGATCANGAAANGATTCAAAGGAAAAGAAANAAACGGAAGGCTTCAAAAATTGACNCT



ATGGGAACCTACCATNGGTTGACACTGAAATCTTTAACCGGAGGGGTCATNTTGCCCGGAAGACT  
CANACTNCTTGACGCCCAAGCCATGNCCTTGCGNNAACACNCTNAGGGGCNATTTCNACACAC  
TTGGCGGGCGTTTCTANTGNNTCCNGGCTCCATNCCAACTTGGC

SEQ ID NO: 705 ACTCCTTGGCTCTTTTGTAGTCTCCACTTTTACTCACTGTCTCTGGGATTTTTAA  
GATCCTTTTCTTCTCTTATAAATCATCCTCTTAATGAAAATTAGCCTAACAAAAGTTGGAGACTGG  
AATCCTACTTTGAGCCACTGACTTGAAATAACTCTTTTGGCAAGTTGCCTGACATCCTGTCTTACCA  
AGGTGGCATATTTGCATTTTTACTGCTTAAACATTTTTTTTTTTTACCATCTTTATCCAAATTTATC  
ATATTGATGGTAGGACTAACAGGCTTTTGTAGAGCTGGCTTTAACTTTGAGTCTCAAGCTACAATG  
CTGTTGGGCAGCCTGGTCTTCCACGTGAGGGTTTAACTTTGTTTATTTGCCTCCAGTTATTCCAAA  
ATGCTTATTAAATGAAAGGCCAGGAACATGTTTATTTTANTCACCTTTGCTTTTTAACAAATTTGT  
TTGTAATCAATGAGTAATTCATGATGAATATTTTTGACTAATGGATAGCCCGAAGGCCAAAGCTT  
TTAATTCTAATAGGNAATGTTCTTTTNNCTNATGGAAACAATGAGAATACTTTGTGCNTTTAA  
AATTGCACTCCGATNTATGCTNGNGGNTTTANTCACATAAGCACAATTNTGNGGTTTATTTTA

SEQ ID NO: 706 ACAGTTCACATTAATATTCACATCCCACTTTCTGCTACTTCTGTGCTAGCTATACT  
CTGGAGAAAAATACAAAGAATGGACTTACACGCTGGGCCAAGGAAATAGAAAATGGTGTATTTT  
GATTAATTGACAAGTTAAAGATGAAGATTGTGACCTATTAGAAGGACAGAAAAATCTTCTAGAG  
GAAATACTCAAGCAACTAGTCATTGTTTTGATGTCAGAGTGCTAACGCAGTTGCTCCTGAATTCAG  
ACCACAGATCCACAGCCACAGTCCAGATATGTAGCGTTCTGTAAACCTTAAGGGTGCTGTGAAA  
TGCAGAGCTTATATCCACAGCAGTAAACCCAAAGTTAAAGATGCTGTGCAGGCAGTAAAGAGGGA  
TATATTGAACACAGTTGCTGATCGTTGTGAAATGCTATTTGAGGATCTGCTTTTGAATGAAATTCC  
AGAAAAAAAAGATTCTGAAAAAGAGTTCCACCGTCTCCTTATCGAGTCTTTGTTCCCTTCTCTG  
GATCCACTGTAATGTTGTGTGATTATAAATTTGGACGATGAGTCANCTGNAAAAATAAGGGACCA  
TTTTATGGGAGATGTTGGATCCCAACAATTCAAATAGAAGATTNGGAANTTGCNNANGAAACAAA  
CANNCTTTG

SEQ ID NO: 707 ACTTGCTTACAGGAAGAGTAATTCCTAGCAAAGGTCATTAGCTCCTAAGGC  
ACTGAGTCAAAGTGACAGCCCTGAAGGAAATTGCACTCCAGCCCTCCTCCAGGATGTCTAATAAG  
ATGGGAAACTTGGATGCCAGCCATTTTGGTGACCTGAGAGTCTAACTACTCCAGTTAGACCTAAG  
GGCACAATGCAGAATTCATGACCTTGTAGTTGTGGCAGGGTCTAGGAAGTCTTCTCCTCCCAAGT  
AAAAAATATTCTCTTGCCATTCTGAAATTCACATTCATATAATGGCTGTGCAATACATGCTTCTC  
ANTAAGAAAATTAAGTGCATGTTTACTGTNTCTGATCACATCAAANTTTTTATGTTTAAAAAAN  
NCTCATTNTGGATTGANTCCAAACNTAGCTCTAATAANANAANNANNGCCCTTNTGGGGAGACTN  
TATTCTTCATTTTTATTG

SEQ ID NO: 708 ACTTTTTTTTTTTTTTTTTTTTTTNGGTATCTATCTAATAAAAGTTTATTTGTGTAT  
GTGCAATGCATAACTCTATCTTANATATGAATCCTAACAGGATGAAAATACTTTCTTGCAACTACT  
TTATGCTTATGAAAGGTGTGAACCTTGCAATGTCCTCCTGTCTTAAACCCAAGTTGACAGTGCCCTC  
TCAAAACTTTTCATAAATAATGACCTAATTTCAATTTAAAAAATGGTTTCAGCAAAATAGAAAAATAG  
AAAGTCCGTTATTTGTCCATTTGTAATATGAGAAAAAAGATGATNCATTCCTCTACAGAAAA  
AGTGGGTTTANAGAACAGTTCTGGTAATATTTACATGGTAAAGTATCAAAAGATCTAATGAGCA  
GCCCCCTTGCTCAGGGAAAGACAGTGATTCAATGTGTTTCTCTTCCGAATTGCTTAATAACTCCA  
AGTGATTTTCAAATTGGGGGGCAAATCTGGCACTGAGTTCATCAAATTTAAATCTATCTTGAATA  
GGGACATTATAGAAATCAAAGAACATCTNTACCCTGCACATNTGGTGGAGTCTGGAGTTANGGGA  
CTACATGCCCAAGTCATCAGCTAAATGANCCAAAATGAACTTTAACCACTNTCTTCANGGAGAT  
GTCT

SEQ ID NO: 709 ACGCGGGAAGCATATGTTACTTACCTTGTTATTAATAATTTCTTGAAAAGCAA  
ATTTAATGGTTTAATTTTATGTGGACGTATGTTAAATTATCCAACCTACCCTATTGTTAAGCATTG  
GTTTTAAAAATTTTATGCTAATATAAATGCTCAAGTAATTTAAATATTGAAAGCATCCCTGTTGGT  
ATAAATTTCTGAGTAAATGCATTGGATCAGTTGGACTTTGAACGCCTTTGAAATGGCTTTGCTAAA  
ATGCTCCCGCCACAAAGTTGTAGGAAATGGGAAGAGGAGTCAACTAGAGGCAAGGGAGTTGAGA  
GAGCTGCAACTGTAAAGGGCAAGAACAGGCAGAGGTAAAAAGATGATGGAAGGTGTGGTGACTA  
AGGGCCACGTTTATGGGTGAAATTTGAGATTGTAGGCCAACTGTATTTTCAAGCTTCTGAACTTA  
GGCAAAATATTCATCGCAAAGTCTCTAGCGTCATATTTTCTCACCCAAATTACGTTTCCCGAGATT  
ATTTATATATAGTTGGTCTATCTCTGCAGTCTTGAAGGTGAAGTTGTGTGTTACTAGGCTGTGTTT  
TGGGATGTCANCAAGTGGCCTGAAATGAGTTGTGCAATAAATGTTAAGTTGAAACCTC

SEQ ID NO: 710 ACTATACTCAGTATTTAAATATGTCCCAGTATAGAANCATAACTTCAATTAAT  
TTGCTNACACTAACTTCTTAAAACTTACAAATATTCAAACAAAGGGAAAAAATNTAGATAGCC  
AAGATTTAAAANGCCATTGCCTTGAACATGGAGCCTCCATAGCAAAANAGGANATATACTTGTAT



ACTCATTATATTTTATTTANNACCCATGTCAGGAAACACACACTCTCTTATNCNNAATATNNTTGG  
NANTTTAATNTCATCENNACTCGAAAACCCAATNTCANCTNTATGGCAATGNATATTCAAATTNA  
TTTACCAANNTNCCNTATTNTANTANAANTCANNAGCTTNAAATATT

SEQ ID NO: 711 ACATGGCCTTTTAAAAACCGGAGACAACTGGAAATTCATTGGACCTGATCAA  
CATCGTAATTTCTATTATTCCAAGTTCTTCGATTGATTTGTCTAATGGAACAAATTGATGTTACCT  
TGAAGTGGTATGANGACCTGATCCTTCANCTCTTTTCCCACTTCCAAACAATGNNTCATTCTTC  
AAGCNTTGGATGTGGCCCATCCGGTNAAAAATGNNTCCTAAAATTTGGGAANATANTAANAATTGG  
GNATTCCCTCCGCAAGTGACCTGAAAAAAAATCCNNTTNCTCATGGCAAGGGACAAGCNCAC  
NANACTTCAGGTGGATTGGCTGACTGNNCTGCTAATATCAANTCTGCGNTTGAAAAGCAACCCN  
TCAAAACAAAACGTNTAAGGATNGGCCNCCCCTTTTAACTGNTTACTNTCCTTTTAAAGGGCTG  
GNGAAACCC

SEQ ID NO: 712 ACGCGGGGACCTGGGATAACGGCGGCGAGCGGACGGCTGCATTACGGGGT  
CTCCCGGAGGGCCAGAGTCGTGGCTTACAGAAGAGACGAAATGTGGTCTGAGGGACGATATGAAT  
ATGAAAGAATTCCGAGAGAACGAGCACCTCCTCGAAGTCATCCAGTGATGGCTACAATAGACTA  
GTTAATATTGTGCCAAAGAAACCACCTGCTAGACAGACCTGGTGAAGGAAGCTACAATAGATA  
TTACAGTCATGTTGATTACCGAGACTATGACGAGGGCCGCACTTTTCTCATGATCGAAGAAGTGG  
TCCACCTCACAGAGGAGATGAATCTGGTTATAGATGGACAAGAGACGATCATTTCTGCAAGCAGGC  
AACCTGAATACAGGGACATGAGAGATGGCTTTAGAAGAAAAAGTTTCTACTCTTCCCATTTATGCG  
AGAGAGCGGTCTCCTTATAAAAGGGACAATACTTTTTCAGANAATCACCTGTTGGCCGAAAGGA  
TTCTCCACACANCANTATCTGGTTCCAGTGTCTAGTAGCAAGAANCTACTCTCCAAAAGGGAGCAA  
ATCATCTCTTCCATCATTTCTTNAACATANAAAGTCCCGTCGTNCTGGTGCCTCTACAAACGGCAN  
AATTGAAGGGAATTCCT

SEQ ID NO: 713 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNAANAACCTTTTTTATTCATCATCTA  
ACCAACANAGGTGGTTGGCTCNAACTCAAACCTAAAATGGCCTCAAAAGGCCACCTNGTTACNAC  
ATGACAGGGCAAAACCANAAGTAGGGACAGAGTTTACCCTCAGTTCTCTGCAAAAAAACCAAGC  
NTNTNTTTACACACAGGTGCCTNATTAATAAACTGATTGGCAATGTTCCACCAGCACANACCCANA  
GTGTGCAAAAATCCGNGGGGGCTCTGTATATNTGTAATTCAAANAATCCTGCNATTTCTTANCCAN  
AAACAANCTCTTGTTTTTTGGGANGAGGGTGATCAAATTGTTTTCTNTAATNATNGCANCAATTTGT  
CAANTCTNNAANTTTTNAAGGGGA

SEQ ID NO: 714 ACTTTCATAATTTGCTCCTGCTATCTAAAAGGCAGAGCCAGGTATACAGGAT  
GGAACATGAAAGCGGACTAGGAGCGTGACCACTGAAGCACAGCATCACAGGGAGACAGGCCTCT  
GGATAACTGCGGGCGGGCCTNACT

SEQ ID NO: 715 ACGCGGGGTGAGAAGGAGAAGGAGCGGCTGGGAGGCGGTTTGGGAGTGCGG  
GGTGGTAACAGCACACGAGAGCGGCTGCTGTCTGCGCTTGAGGACTGGAGGTCCTGTCTAGGGA  
ACTTATANAAATGCTGGCAATTTCAANAAACCAAAAGTTGTTACAGGCTGGAGAGGAAAACAGG  
TCCTGGAGTTGTTAATTCACCGAGATGGGGAATTTCAAGAACTAATGAAATTGGCACTTAATCAGG  
GAAAAATTCATCATGAAATGCAATGTTTTAGAAAAAGANGTTAGAAGAGAGACANTGATATTCAC  
AGCTACAAANACANCTTANGGANCATAACAAATNCTGGCAACATCTTNTTTACCANNCAANGAGA  
AACTCNGTCATAAAAAAAGNAAGAAAAAGGTCTTTTNTCCTTTNAAAAAANNATTAATATNCNCN  
TAGGATCNNGNGCAATTTNTNCTNTTTGTNCCTCNCTTNNCCTTGGTTCNAGGGGAACCCCNNGGA  
TAACCTTTCCNAANTNAATTTANAAATNTAAAAANTGGGTTTANTGGTNNTAATNTAAACAATTCT  
TTCNNTTTTGGGGTNATANTGGCCTTTTNCNAAAAAANTCCCTTCTNTNNNNGAATATTTNCCA  
AATTCCTT

SEQ ID NO: 716 ACTAAATCTAGTAAAGACATTTTCATACACACCAAGGGGAAAAATAGGTAGC  
ATTACAGAAATTTTGTATGCAAGAATATATTTTTCTTATTTTTTTATCATGGTCTACGAGAACAT  
TGTTTGAGATGCCAAATATTTTCAATTAAGTTGATGTTTCTTTTCTTTCTTTTCTTTTCTTTTCTGTA  
GACAGGGTTTCGCCATGCTGCCCAGGCTGGTCTCAAAGTCCTGGCCTCAAGCGATCCACCCACCTT  
GGCCTCCCAAAGTGCTGGGATTATAGGCATGAGCCACCTTGCCTGAGCTGGTATTTCTTGAGTATC  
TACTTTGCAGAAAACAACCAATGAGTCCAGTGGCATAGCCTTTCTCTTTCAGTATTTTGCAAAAG  
TTGTCTCATTTGTTGGGAAGACCTCCAGATGCTCCGGTCCACTGAAGACACGGTNACCAATGCTGG  
AAACCATCCCTGATCGCACAGGGTTTCTGCCCGTGAGGAANGGCNGCTCTGCTTGGGGNTGCC  
AAAANATGCGGCTNAAATGNGNTGGTCNACT

SEQ ID NO: 717 ACGCGGGGGCAGATCAGGGATCGCGATTGCGAATCCTCCGCTGAGGTGATTT  
GGATATCCCTAGAACGTTGAGGGCACGAGTCGGGTCTGAGACCAGGTCTCAGCCAGCAGAGCC  
ACGTTCTTATGAGCACCGTGGGTTTATTCCATTATCCTNCAACCACTGACCCGAANATGCCCGGCG

CCNTGGGGACTCCNGCTT

SEQ ID NO: 718 GCTTTTTTTTTTTTTTTTTTTTTTTTTTTTACTTTTGGGAGGANATAACCAGTCTC  
TCCCTTCATATATATTCTTTTTTATTTCTTGTATACCTTCCCAAAACAGANACNTTCAACAGTAGT  
TAGAATGGCCATCTCCCAACATTTTAAAAAACTGCNCCCCCAATGGGTGAACAAAGTAAAGAG  
TAGTAACCTANAGTTCAGCTGAGTAAGCCACTGTGGAGCCTTAAGTGGTGAGGTCTTCCAATTTCA  
GAGNGATGTGTCTTCAACTTGTATCATCATTTTAGCGGAAAAACATAATTTAATTTTGGTGAAATG  
AGATTCATNTCGTAACAGGATTAGTAACAGCATTCACTGAATTTACACTCTTTTTTTGTGAACAA  
GTGAAAAAATGAATGGTAGCTAAAAANATCAATGGGATTCCANCTCCAGCTGCANATGGAATGAC  
CATTTACCAGGACAAAAAACCTNTTTTTTGTATTTACAAAGCTAACAGCAGTTTAAATCTGCAG  
CTTAATTCNGACCAACTACCCACACCTTTTNAATGGCAAAAAATACAAAAAGTGTTCAAAANTN  
TTTCAAAAAAGNGCATAGTCTANGTGCTTGTAAANTAACCCTGAAAAATTNTTGCTAGAGGGA  
GTTTANC

SEQ ID NO: 719 ACGCGGGATGGGAATGAGGNTCTACCACTCTGGAAAATTCATGCCTGCAGGT  
TTAATTGCAGGTGCCAGTTTGCTGATGGTCGCCAAAGTTGGAGTTAGTATGTTCTACATGACCCNA  
TTAACNAAAGTNTNTTCCACCTNAAACATGATAAAATAATTAATAAANT

SEQ ID NO: 720 ACAAATAAAAGTGATGGTGAGAACCTGGCTCAGGAAATGCAGTAGCAGGCC  
ATATTGCATCCAAAGGAATTACTCACAGCTGTGCTGTGTGCATTCTCTGTGGGCCTAGCAGGGAAG  
GGGACAGCCCTGTGGCAATGGGCATGACACGGATGCTCCTGNAATGCAGTCTCAGTGACAAGTTG  
TGTGTATCCAGGAAGAAGCGGGTATGAAGTGATTATCGTCCCACTTTGTGGTAACTATCTTCC  
TCATCCTTCTTTGGGGTCANTCCTNGTGCTTTTTATCANAGACAAAAGAACTTAACACCNGCGT

SEQ ID NO: 721 ACTTTTTTTTTTTTTTCTTTTGATTTCTCAGGACCTTAGAGGGAAAAACAAAC  
AGTAGCAGCTAATATTCTCAAGTATATTGCTGCTTANAAAGATCCTCAGGAACAATTAGCAGCAAT  
AAGCAAGCCTTTGAAAAGATCTGAATTTCTTTTCTGAAATATTTACGATACACAGGTGCTTTTTAT  
CTGAAATCTGTTGGGTCCTCCTTTTAGGCAGTCTCTGTGGGCAANAGAGTGGGACTTGCNAGGTGG  
ACAGCTGTGNGGGATCCTGGGGAANGGAGTTTNAANAANGGGTGGCTCAGGGCNGTNTAAAAA  
NCCNTTTTGGNANGGATTNGTATTGAAAAAATNAAAACCCTTTGGTAAGGCN

SEQ ID NO: 722 ACCGGAAGAAGCAGCTGGCAAAGCAGCTCCCTGCACATGACCAGGACCCTTC  
AAAGTGCCATGAGTTGTCTCCAGAGAGGTGAAGGAGATGGAGCAGTTTGTGAAGAAATATAAGA  
GCGAAGCTCTGGGAGTAGGAGATGTCAAACCTCCCTGTGAGATGGATGCCCAAGGCCCAACAA  
ATGAACATTCCTGGAGGGGATAGAAGCACCCACANCANTGGGGGCCATGGAGGACAAATCTGCT  
GAGCACAAAAGAACTCAATATTCTGCTATTGCTGCAAACCTGAGCATGAAAGAAGGTAACCCAGC  
CATNTATGCCTAAAGGGCTTGGCTATTATAAACTGTGGCACN

SEQ ID NO: 723 ACAGCGTTTACAATGCTGGTATTAATCAGCTACATATTTTGAACATCTACTGT  
TACTGGATACCAAAGAAAGTGAGTTATTTAANAATCTTCCATTCTTGTTATAAGCTTCNTATNATC  
CAGTAACCTTCTCANANGCTTTNCAANAAGCNTAAGTTCTTGCTTTGAGANAATTNTAATTGTNC

SEQ ID NO: 724 ACGCGGGGAGTGCGTGCCGCTCCGCCGACCGAANAGGCTGGACATGACACC  
AGTGGCATATCACGGCCATGGGGTCTCAGCATTCGNTGCTGCTCGCCCCTCCTCCTGCAGGCGAA  
AGCAAGAAGATGACAGGGACGGTTTGCTGGCTGAACTAGAGCANGAAGAAGCCNTTGNTCATTN  
TATATGNTGAATNCCCGGGAGA

SEQ ID NO: 725 ACCTCATCGGTATCCAAGGCCCGGACTATGTTCTTGTGCGCTCCGACCGGGTG  
GCCGCCAGCAATATTGTCCAGATGAAGGACGATCATGACAAGATGTTTAAGATGAGTGAAAAGAT  
ATCTCCATTCTCTTAAAGTAGTGGTTCTTTTTGCCCTTAAACTTAAATTTTTTAAATTAATTAACCTC  
CGCCTCCCGGGCCCAAGTGATCCTCCCATCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCACAC  
CACCAATGCCAGTTAGTTTTTGTGTTTTTATAGAGACAGGGTCTCACCATGTCATTACAGTTGGT  
CTTGAACCTCTGGGCTCAAGCAGTCTGCCTGCCTTGGCTTCCAGTGCTGGGATTACAGGCGTGAG  
CCACCGTGCCCGGCTAAAAAGTATTTTAAGTTCTGCATATTGCTTATTTCACTTAACACTATATTA  
GAGATTGTTTTATATCAATACATATAGATATGCTTATTCTTGTGACAGTTGCATAATTTTCCATTA  
AATTGATGTATCATGGGCAGTTAACCAAGTACTCGTTTTACTCTTANCATAACTTTGGGGGAACA  
AATGTGGATGTTTTGGNGGTTAAAGCTATTAAAAACAGGGGTTTTTGCCCN

SEQ ID NO: 726 ACGCGGGTATAGAGGGCTAACTCAGGCATTGTCTTGTTTATTTGTAGACTGGA  
TTAAAAACAACCTGTCTGTTTTGTGAGTTCCAGCTTCTCGTTTGAATAAAATTAGACCAAAAG  
AAGAAACGTGCTTGTCTGTATACCCGCAGAATGAAGTTACTGTTGTTAAAACTGGATTTTTTCA

TTTACTAGGTTCCGAAGAGTCCAGATGCTTGGTAGATGTTCAATACNTGATTTTTTTTTTAATNGA  
ATGTGTTCATTTTAAATCCTCCTTAACATTTCTAG

SEQ ID NO: 727 ACTTACATGTGTGAACACATATAAAGTGCCAGGTTTACAGACCCTGGCTCAA  
GGACAGTCTAGGATGGGAAAGGAGGTAGGGCGAGAAGAAGCACATATTTTCTCCCTGGTGCTTCA  
GCCTCACCCCTATCCAAGGGACAGACATATGGGGTGTGAGAAACCCATCCCCAGGTCCCAGCCTTC  
AGGACTGGAGTCCTTTTGAGTCTGGTGGAGTCACAGATCCAGTCTTTGGGGGACACTGGGTCTGTC  
TCCTTTTGAAAGCCCTGGAAAGGTGGGAGGTAAGAAGTAAAGGGAGACAGGTCCCTGCTAGAAG  
AACTTGACGCCTCGGCCATCGCTGACAGTGATGATCTCGGCCTTGTGCTCCTGCTGTAGGGCCTGC  
AGAGCCCGCAGTAGAGTGGCTTCATCCAGCCCGTGAAC'TCCTCATCCTCTGTGTCTTCCCCATTA  
GTCAGTTCATACAGGGTAAAGACGGAGTTGTTCTGCCACTCCTGGAAACCCACTGATAGATGAGTT  
TCCCCATTCTTCTGGCCTCCGCCACATGATCAGGAAGCTGGACTTGCTCTTATCCAACCACTCGA  
NGTTCCCTTTCTTCTTAATTCTCTAATAACAATCTGGATCGACTCCACAGGAAGCTTC

SEQ ID NO: 728 ACCACTGTATTGATTAGCCTGTATGTAGCAGGGCTCCCTTCATTGCATCTGAG  
GACTTGTCTTTCTTTTCTTTATTTTAAATCCTCTTAGTTTTTAAATATATTGCCTAGAGACTCAGTTAC  
TACCCAGTTTGTGGTTTTTTGGGAGAAATGTAAGTGGACAGTTAGCTTTTCAATTAAGACACT  
TAACCCAA

SEQ ID NO: 729 ACTTCAGGATTAGGAATTTGGGTTTGTCTAGATGTATTCTCTGGTGAGGGTG  
GCTGGGATATACCTGACCCACCATCTTCAGAAGGACCCATGTCAGGTCTGACCATTGGGAGCAAA  
GCCATGTTACACTGACCTAATGCAGAGTATGGAAGCATTTGGGCTGGTTATACATTTCTGTTTCTT  
AGATTTATCTCCGCTCTGTAGGCATGGACAACCTTTAATCAGAGCATCTAGAGTGGCCTCTTGT  
TTATCCTGAAGATACTGATGGGTCTGTTTTCTGTTAGTCTGTTTTGTAATATTCTTTCCCTTCCTT  
CATGGGGAGGCTTAGTTTGTCCAGTCTTCCATGCCCTTCTATCCCAGATTACCTAAATGTTCCCTT  
CTCAGGAATTCTGTCTCATCAGTTCTTCACAGTGAGAAAAGAGGCTAGATGATGGTGTGGGGGGTT  
GGAGTTTTCTTCTAATACCGAGGGTTCCTGGCTGTGAGGAAACAGCCACATGTTTCGTCATGATTGA  
GCTGTGAAGTCTTCTTGACCTGTTGCTGAAAATAAAGTTAATTTGTTTGAGGCNTCTCTCTTAAGT  
AGGTGGAAACTATTGAAGTTANCTNACAATCACANCATAGGTTCTGATCCTTGAAAGGGGGTTG

SEQ ID NO: 730 ACCTCCTNTTCTCTTCTATTTTTAGGAANAAGTTATAACAAGTTTTAAATATCT  
CAATTCNAAAAACAATAGGCTTTTAAAAAATAAGACTTGNTTACCAGAAACAAGAAATATGTAG  
TTACATAACCATTTTCATATCACTACTCATTTCATNATTTACCAATTCATCTTTGATGCAACTNGN  
GAAACNTTTAANCAGTCACTAGACACCTGTTTTAGAATCTGAAGAAATTATTATCCACCAACAGG  
AATCTAATGATATATATATTTGCATATATTCAAAAATTNCATGAGGGGAAAAAGGTNNTAAANTATT  
CCTAGTTNTTTTATACCANATATNGATATTCCTAGNANNAACAACCAAGGNNTTNAATNTTTTTTC  
NNAGNTGAAACTGNTGACTTATATATGAAGCTTTTTTCCAGNACTTTTTTNCGGGANATTAAAAAA  
TNNATTTTATTTCCCTGCTTGGGCTAAATTAATCAATTAATTTATTANCAAGAGAATTTCNTNAATA  
AGGANAAATTNATTTCCCTTANNANAACCANTAATATNNTTTTTTTAAAAAATCTTTTAAAAATCN  
NTNAAAAATCCTNANCNTTTAANCNTATNGCATNAATNCCCAACTNGNCNTAANAAAAATAA

SEQ ID NO: 731 ACCTCCTCTTCTCTTCTATTTTTAGGAAGAAGTTATAACAAGTTTTAAATATCT  
CAATTCGAAAAACAATAGGCTTTTAAAAAATAAGACTTGATTACCAGAAACAAGTAATATGTAG  
TTACATAACCATTTTCATATCACTACTCATTTCATTATTTACCAATTCATCTTTGATGCAACTGG  
AAACAATTAAGCAGTCACTAGACACCTGTTTTAGAATCTGAAGAAATTATTATCCACCACAGGAA  
TCTAATGATATATATATTTGCATATATTTAAATTTTCATGAGGGGAAAAGGTAATAAACTATTCTAG  
TTATTTTATACCAAATATTGATATTCCTAGTCAAAACAACAAGGAATTAAGATCTTTCTCCAGGTG  
AACTGCTGACTATATAGAAGCTATTTCCAGCACTTTCTTCTGGGGATTAAAAATGATTTATTTCCCT  
GCTGGCTAAAAATGATAAATTATATTACAAGTAAGTCCTTCAGAAGTGAATATTAATTCCTTAAG  
AGACCATAAAATGTATTTTTAAAAAATCTCTTAAAAAAGTAAAAACAATCCTTCTCTCTAAACCTA  
TGGCTTAATCCCTACTTGACCAAAGAAAAAAGAAAGT

SEQ ID NO: 732 ACTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTAGTTAATAAAATCAATACAAAT  
CTTTTATTAAGATCTACTCATACCATGGCTGAAATCATCTATTATGTTGCTAGTTAGCCTCTCTT  
CTATAGTTGGGTAATGTTGCCTTGCNACTGNTTTNCCATCTCTCCCAAGTGAAAANAACACTTTTT  
AAAAAAATTANTTGCTCCAANTTTTINAGGCCANGGGAGGCTCTCCCATNNCNTCCTTAAATA  
GTCCNCTNCAGGAAGGGTGATCTTGNGTATAAAATTTTCATACTTAANTTT

SEQ ID NO: 733 ACCAGGCTGGCGACAGGTGCTACCAGGAGTGGGCTGAGGGGAGAAAAACTA  
TCTCCCACTCTTTTGGCCAGGCAATGTCAACGACTTCCACATTCCCTGGCCCACTTCTGAGCAA  
CCCCAGGTTTCGGCTCTGTATAAGGACCTCCCCTCCCAACCCCAACCCAGAGTGCAGTGCAATC  
AACCAACAATTTACTGGTGGAATGGCAATCAAAGGAAACAGTTAAACACCAACAATTTCTTAAA

GCCAAAAAATATTTTTCATGGAGTTGAACATTTTTCGAGTGTGTTTTTTC AAGTGTAAGCAGT  
GACATTTTGTTC AAACAGAAGCAGCATCTAGGAATTCTGGCACTTGGGTTCTAGGGGGTTACAGGT  
ATGCATCATGGATTCTTCTCCCTCGTATTTAAAAAGGCCTCGTGTTCCTAGAGTTCATACCA  
ACACCTGCTAGCTCTCCCTCTAGCGGACAGTGGGTGGGCAGCCAGCCTCCTGGTTAGATTGGGCA  
ATGCCAANCAAAACATNCCTNATTCACCTGCTTGGGCTTGCTTCTGATT CANANGTAAANTCAANT  
GCANANAAAGAACTTACAAAANCNCACCACCAAAGGNANCTNNACGGGGACCCTGNCNNANTG  
ANTTGNTG

SEQ ID NO: 734 CAGCAGAGATATATGCCTATCGAGAAGAACAGGATTTTGAATTGAGATAGT  
GAAAGTGAAAGCAATTGGAAGACAAAGGTTCAAAGTCCTTGAGCTAAGAACACAGTCANATGGG  
AATCCANCNAGCTAAAGTGCAAATTCCTCCCGGAATGTGTGTTGCCTTCAACCATGTCTGCAAGT  
CAAATTNGAATCCCTCAATAANGTTGCCAGATATTTNCCTTCAAACCTTGCTCAAGAGAAANACC  
CANTGTTNTTTTAAATGGGTGGCAGAAANTNCNTTAAGAGNAAGNTTCATTGTGCAAATTCCTNAC  
TTTCATGGCCTTCNCTGGCTGNTATTCCCTTATATGATGCTGNNACCCTTTATGGGACNGAATTCAA  
GAAAC

SEQ ID NO: 735 ACATTGAGACAGAGCTAAAGAAGAGGAAAGGGATCGTGGAACATGAGGAAC  
AGAAAGTTAAGCCAAAGAATGCAGAGGACTGTCTTTATGAACTTCCAGAAAACATCCGTGTTTCC  
TCAGCAAAGAAGACCGAGGAGATGCTTTCCAACCAGATGCTGAGTGGCATTCTGAGGTGGACCT  
GGGCATCGATGCTAAATAAAAAATATCATTTCCACGGAGGATGCCAAGGCCCGTCTGTGGCAG  
AGCAGCANAACAAGAAGAAAGACAGCGAGACCTCCTTCGTGCCTACCAACATGGCTGAATTAT  
TGTGCAGCACACAGATTTTATCATTTGAGGAGCTCAACGCGCCATACGGAGAAACAAAGAAGAG  
CCCAAGGCCCGGCCCTTGAGAGTAGGTGACACGGAGAAGCCAGAGCCTGANCGTCCCTCCTANC  
CGCAAGCGTCTGCTAACGAGAAGGTAAGTGGATGACTNNTNNTTTTGAAGTTCAAGAAAAAT  
GACTAGGCNGTNCCTNCCCGGGCCGCTCTNNGGCCNANGTCCAAAAAAACCTTTNNTTATTAN  
GANTGNATTCTTTACNGGCNTAANTGCTATCCNGTTCCAANTCAAACCAAGTTTCTCCC

SEQ ID NO: 736 ACGCGGGATAGACGGAAATGGAGAGCTGGATTTCTCCACTTTTCTGACCATT  
ATGCACATGCAAATAAAACAAGAAGACCCAAAGAAATTTCTTAGCCATGTTGATGGTGGA  
CAAGGAGAAGAAAGGTTACGTCATGGCGTCCGACCTGCGGTCAAAACTCACGAGTCTGGGGGAG  
AAGCTCACCCACAAGGAAGTGGATGATCTCTTTCAGGGAAGCAGATATCGAACCCAAATGGCAAAGT  
GAAGTATGATGAATTTATCCACAAGATCACCTTCTGACGGGACTATTGAAGGAGGAGAATGG  
GAGAGCCTCCCTGGGCCTGAAAACCTTGGAGTAATTAATTTTAAAAAGTGTTCTTTTCACTTG  
GGAGAGATGGCAAACACAGTGGCAAGACAACATTACCCAACTATAGAAGAGAGGCTAACTAGCA  
ACAATAATAGATGATTTACCATGGTATGAGTAGATCTTTAATAAAAGATTTGTATTGATTTTNTT

SEQ ID NO: 737 ACATAAACTTCAAAGAGATGCTGTAGAGGATTGGACTGCAGTTTTTCTCATA  
GCCAAACAGCTGGAAGCCAGTTCCCAAGACCAAACCTGCATGCCCCAATCGCCAAGGTAATTTA  
TTCTTATTACTTGATGTCCTAAAGCTTCTTTGAGATTIGCTATAAAATTTCTATGATGGTAGAACG  
CAAATGTCTCAACATGAAATTTTGGCAACATTAGGTGAACTGAATTC AACCAATCTTCTCTG  
GGGAAGTCCAGAGAAAAGTTCACCTTTTAAATCCATATTTTGAGCCATCTTCAATTACTTGTTGAG  
CACTGTCTTTGTTAAGAGCTCTCTGTTTATTTGAAATTTACAGTCCTTTGACCAGT

SEQ ID NO: 738 ACCAAGTGAGTGGGAATACATATTCTAGTTAAAGCATTGTGTCTAGCTACAC  
ACCGCTAACAAAGTTACTTAGTTATCAATGTAGGATTCTTAAGGAGCTTTAAGCTAAGGAAACCTT  
TAGTGACTTANCTTATTTGTATCTTTTCACTTAGGAAGATTTTGAGGT

SEQ ID NO: 739 ACCTTCACCTGCTCCAGTGTATGANAGCCTCCAGCAACATATAGAAAAGCACA  
ATGAACTGAAACCTTACAAATGCCAGCTCTGCTACTATGAGACCAAGCACACGGAGGAACTGGAC  
AGCCACCTTCGGGATGAGCATAAGGTAAGCCGTAACCTTTGAGCTGGTTGGACGGGTAACTTGGA  
TCAGCTGGAACAGATGAAGGAGAAAAATGGAGAGCTCCAGCAGCGATGATGAGGACAAGGAAGAA  
GAAATGAGCAGCAAGGCTGAAGACAGAGAGCTGATGAGATTTTCTGACCACGGGGCTGCTCTTAA  
CACTGAGAAGCGTTTCCATGTGAATTTGTGGACGGGCGTTTNAACAAGGGCTNTNAGTGGGAA  
AGACATGTGCTNANACACGGGATGGCATNTAATGANACCAAGCNNGGTGAGCANATANNAAT  
CNCCANAAAGAGATCATGGAGAACCTGTTTAAATGCCCTTCNTNTANGGAAANNGAAGATTGA  
CTANGCCCTTG

SEQ ID NO: 740 ACTTTTTTTTTTTTTTTTTTTTTTTTAAACANAAAGGTATAAAGTTTATTAACATCT  
TTAAAAAAGATGGGCCGGGCTGNGGCTCACNCCTGTGATCCCAGCACTTT  
GGGAGGCCAAGGCGGTGGATCCTTTGAGGTGAGGAGTTTGAACACAGCCTGGCCAACNTGGNGA  
AACCCCATNTTTACTAAAAATACAAAAAANTTANCCAGGCNTGGNGTCGCACACCTGTAGTCCCC  
AGCTACTCGGGAGGCTGAGGCAAAAAATCGCTTGAACCTGGGAGGNGGAGGTTNAGTGAACN

AANATCGNCCGTTGACTCCAACCTGGGCAAAAAGCAAAAANTCCATCTAAAAAAAAAAAAAAAA  
ANGATGGGCCNNACCNCNGCTACTGAATATTGATATAGNTCCTTTTNNACAAGTT

SEQ ID NO: 741 GGTACGCGGGTGAAGTATTGCTAATATTACCGNGGTTTATGAACTATGTTTCAG  
AATTGAAGAAAAATCCTAACCTTCAGTTAGAGGTTAGTGACGGGGTTCAGGACACCCTACACAAAA  
TACAGCACTTTGACATATTGAATATTTAAGCTGAAGGCATTTGAGGAAATTGCAGAAGCAGGAA  
GGTGA CTCTGACCTTCTGCCTGCTGTTCTCCCCAGAAGCAGCCATAAAACCTGGGAAGGATTTTCT  
GACCTTCCCCTGAAGTAGATCATAAGACTGTCATGTAAGAGGTGCTCTCCTGGCACCCAGAGAAA  
AGGAGCATCCTTACCTCAAAAGCACAGGGACACAAAGAGGAATCTAAACAAACAGGCCTCTCAA  
TTTTCCCCCAGTTATTACATTTAGCTTGGTCACACTTTTGCCTATGACATTTCTACATCACTGGCTTG  
NTCTTCATCAAACCTACTATNAAAAACATTCAAGTCAACTGTTNTTTGGGCCTTTTCTTATGG  
AGCCCCNTNNGGTNGGGTAAACTTTATTNAANAAANNNGGCNTNCTTAAAAAAA

SEQ ID NO: 742 GGTACTGTTGGTTAAATGACAATTTATGTGGATTTTGCNTGTAAATACACAGTG  
AGACACAGTAATTTTATCTAAATTACAGTGCAGTTTAGTTAATCTATTAATACTGACTCAGTGTCT  
GCCTTTAAATATAAATGATATGTTGAAAACCTTAAGGAAGCAAATGCTACATATATGCAATATAAA  
ATAGTAATGTGATGCTGATGCTGTTAACCAGGGCAGAATAAATAAGCAAAATGCCAAAAGGG  
GTCTTAATTGAAATGAAAATTTAATTTGTTTTTAAATATTGTTTATCTTTATTTTGTGGTAA  
TATAGTAAGTTTTTTTAGAAGACAATTTTCATAACTTGATAAATTATAGTTTGTGTTAAGAAAAG  
TTGCTCTTAAAAAGATGTAAATAGATGACAAACCGATGTAAATAATTTGTNAAGAAGCTTCAA  
ANTGTTTATACCGTGGAACACACCTACNTGAAAAGCAGAAAATCGGTGCCTGGTTTGCTTCTTTT  
CCNCNTTATTTTTNNGGATTGGGGGCNATTTCCCATNCAAAAAATGGGGGCCNANACCNGTTNT

SEQ ID NO: 743 ACTGGGATTATATAGGCATGAGCCACTGAGCCTGGCCCANAAAGCGTTTTTCTC  
AAAGGCCCTCAGTGAGATAAATTAGATTTGGCATCTCCTGTCCTGGGCCAGGGATCTCTCTACAAG  
AGCCCCTGCCCTCTGTTGGAGGCACAGTTTGAATAAGGAGGAGGAGGGAGAAGAGAAAAATGT  
AAAGGAGGGAGATCTTTCCAGGCCGCACCATTTCTGTCACTCACATGGACCCAAGATAAAAGAA  
TGGCCAAACCTCACAACCCCTGATGTTTGAAGAGTTCCAAGTTGAAGGGAAACAAAGAAGTGTT  
TGATGGTGCCAGAAAGGGGCTGCTCTCCANAAAGCTAAAATTTAATTTCTTTTTCTCTGANTTCT  
GNACCTTGCCGNGACCACCTN

SEQ ID NO: 744 ACTTTTCTTTCTTGCTGGTAATTTTATGGAGCAGGTTAAGAAGGCTGCTCTGT  
GTTAGGATAAACTGTATACCAATAATGTTGACAACCTGTAATGAGTGTGCAATTTTACTTCTGTAT  
CTTTTCTTCTACCTTGATGCCAGTAATCTATAAGGGATCTTTATAGTTTGAATGTATTTGAATAA  
CTTCAGTATACTTTAGTTCTACTTTTTTATTTGACTCACAACCATTTCTAGGTCTCAAGTATTCCCAT  
GTGTTTAAAGCCTGAAGTCAGTGAGATGAAATTCACATCAAGAATTTGAAGTAACTTGTAAG  
GAAAAATAATATAAAGATACCATTTGGGGCAGTGGCTCACGCCTGTAATCTCAGCACTTTGGGAGG  
CTGAGGTGGAANGATCACTTGAANCCAGAGTTTGANACCANCCTGTGCAACAAANCAAGANCCCN  
CTTTCAAAACTTAAAAAATANCTGGNTGGGGGGTGCTCACCCCNANTTCNCTANTCGGGAAN  
CTGAGGNNNNAAAACCTTNNGCCNNNAGGCCATC

SEQ ID NO: 745 GGTACTGCTATAAACTCATTCTTGTGTGGTGGCTGTGCTATAGAGTCTGTGT  
ATTGCTGTTTCATATTCGGAGTTCTGGTTTTGTTTTCCCTTAAAACCTGTAAACAGTTTTTTGGGGG  
TGGGGGGATTCAGAACTCTGTTTCCCATTCATAGCACCTGACATTATTCAAGTTTTCTAATATC  
TTAAGGTGTATATTTATTTTTTTTATTGGCTTAGTTGTTTTGTTTTGTTTGTGAGATGGAGTT  
TCACTGTTGCCAGGCTGGGGTGCAATGTTGTGATCTTGGCTCACTGCAGCCTCCACCTCCCGAGT  
TCAAATGATTCTCCTGCCTCANCTTCTGAGTAGCTGGGATTACAGGTGCATGCCACCATGCCCGG  
GCTAATTTTATATTTTATTANANACCGGATTTCCCATGTAAACCAGCTNGTCTCGAACCTGACCT  
CAGNGAACCTGCCCCNTGCCTCCAANGNTGGGATAACAGCTGANCCCCATNCCGGCCTTNTG  
GCTTANNTTTTAAATATCCCCCNANN

SEQ ID NO: 746 ACTTTTTTTTTTTTTTTTTTTTNGCCTGGAAATGTTTTTAATANAATTGGTCTAG  
TAATCGTTCAGGATTTCCGTGATGGGCCCTCCCTGTCTGGACACTGCCAACCCACAGCTGGAGGGG  
CACTTAAGGCACGTCAATTTGTGATTAGAATTACACAAAATTTGATTAATATTATAGCTGCAAAAT  
TAACATACACAATTTTCACTCATAAATTTAAATATTTTGGATGAAATCTTTGCTTTTCAATAAAAG  
ATCAATGGTACTTTTTTTTTTTTTTTTTTTTTTIGCGATGGAGTCTTGCTCTGTTGCCCGGGCTGGAG  
TGCAATGTCACAAATCTCGGNTCACTACAACCTTTGCCTCCGGGTCAAGCGATTTTCTGGTTNA  
ACCTNCTGAAAACTNGGACTACCGGGGCCCCCCCCCCTCGGNTAAANTTTGNATTTTAAAGNA  
AAAANNNGAGTTTACAATTGTTGGGNAACTGGNNTTAAANTCNNGACCTAAANGANCCCCCTNN  
CTTGGNNTCCNAANGGCTGGGAAAAANGNTTNCCT

SEQ ID NO: 747 ACTGTGATGTGAAGCTGTTGCTGGTAAAATTAGCATTGGCCTCCTGAAAGGT  
CTCTCAAATGGGATTTTTTCAAAGGTTGCCAAGACTTGGACTCCCATTTTGTGCCAATTGCAGGGA  
TTCATGTTACCAGATGTAATGTGGTAGACTAATGTTGACTTAGGTCTGTGAACTGCAGTATACCAT  
CCTGCAGCTAGCATGAGATTGACGACTGTATCAACTGGAATTACGTCTGCCACAGCCATTGGAGTA  
GCTTTTATGGCCCGAGGAAACCCCTTCCAGTCGCAATAATGATTCCATTAGGTCCATTTATATTAT  
CAACCCAACTGGGAAAGGCTCCTGCCAAGTTGCTCCCAATGGAGGGCCTTATGATGGCAATG  
TTCGGGTTCTGCTCTCTTGTGTCACCACCATTNTTCCAANGNCTTGGTGTAAGTATAAATATTTG  
GGCCCATCTCTGATCAACCTTGGNGTAATCTCGNCATNAAAACCATNGCTAACCCCTCAAGGGG  
AATNAAGGATTTTTTTTGGCTCCCCCGGGCNCGGGTANANAACCTTCTCCAAGNGCTTTANG

SEQ ID NO: 748 CGAGGTACGGGGGTCTTGAGCGCAGAAACACTTACTTTTCCCCCTACCCTGCT  
CCTCCTCCTCCACAGCCGTCTTTCTCTTTGCCTCAGCCACTTCCTTCCTTGGCCTCACCCCTCCCCAGT  
GCACTGAAGAAGGTAACCGGGTCCAGACCCACGCGGCCAGTTCTCCGGCGGGAAGGAAACC  
GCGCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTCATCGTT  
TGGGTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCTCTTCCGTGATGATAAAT  
GTGCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACA  
TATCCAGGAGAGCTGCTTCTACAAGGTGTTTCATGATGATGTTGACATTATATTACTGCAAGAATA  
AATGNGGGTTACATATCTTTATGTACCCCGGCTNNATTGGATCCAATACTTGCCAACNGAACCAGT  
TACCCCTANGGATACAACGCAATNCTANTCTAGAATCCATATTAACCATAGGGTTACCAACCTCATG  
GTGGACAGGAATCCNATGGNCAACCCNTTTAAAGGTCCTTGGTCAACAATAAAGNCTACCNGNC  
TTAGTTAAACGNGTAATCNAGGTGGGTNTATTACTTAAATCTCTNACTNCCCNGGCGCTTTAAG  
GGAATTA

SEQ ID NO: 749 ACGCGGGATTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAANNACACGGGGTNAATTACACAGTTCTATTAAACCTNATGCCTTTTC  
ATTACATNTAATTTGAACTCTNAACTTCATGTTACAGAATGCTTTAAAGATGCTTTAATGAAAAGT  
NTTAANAAAAATNTATANATTTGTATGTACAGTTTATACTTCAAAAATCCATATATTTGTATATTAAT  
TTTTTTAGAANCCTCCTAATTGGATAACTAGATGGTATTTAAATGAATGCCCAAAAATNTCTTGT  
CCCCGCTGTTTACCAAAAACATNACCTTTAGCATCACCAAGTTTATAGAGGCACCGGCTGCNANTGA  
CACATGTTTAAACGGCCCGGNCCTTNGCCGNGACCACCTAGGGCNAATTCACCTANTGGNGGGCG  
GTACTANGGATCCCANCTCGGTCCAACCTTTGCNTAACATTGGNATANCTNGTTTCTGTGTGAAAT  
TGGTTTCCGTTCCAATTCCCACAANATTCNANCNCGAACNTAAANTNTAANCCTGGGNNCCTANTA  
AGGACTACTNCTTTATGNGTTNCCTACTTCCCTTTAAATGGGAACTTTTGGCCCTNTTAANANTTG  
CCACCCCGGGAAGCTTTTGNNTNGGGCNTTTNTTTTNTACTATNNN

SEQ ID NO: 750 GGTACGGGGGTCTTGAGCGCAGAAACACTTACTTTTCCCCCTACCCTGCTCCT  
CCTCCTCCACAGCCGTCTTTCTCTTTGCCTCAGCCACTTCCTTCCTTGGCCTCACCCCTCCCCAGTGC  
ACTGAAGAAGGTAACCGGGTCCAGACCCACGCGGCCAGTTCTCCGGCGGGAAGGAAACCCG  
GCAGAGAGGCAGCAATGAATGTGGATCACGAGGTTAACCTCTTAGTGAGGAAATTCATCGTTTG  
GTTTCAAAAAATGCTGATGGAAAGTTAAGCGTGAAATTTGGGGTCTCTTCCGTGATGATAAATGT  
GCCAACCTCTTTGAAGCATTGGTAGGAACTCTTAAAGCTGCAAAACGAAGGAAGATTGTAACATA  
TCCAGGAGAGCTGCTTCTACAAGGNGTTCATGATGATGTTGACATTATATTACTGCAAGAATAATG  
NGGGTTACATATCTTTATGTACCCCGGCTCAATTGATNCCAATACTTGGNCCACGGAACAAGGTA  
NCTAGGGATAACAACGCAATNCTNTTTNTNAAAGTCCATATTAACAATANGGGTTTACNACCC  
CATGTTTGGATCAAGGA

SEQ ID NO: 751 GGTACGCGGGAACCTTTGTAAGATGCAAAGAGGTTGGATCAAGTTTAAATGAC  
TGTGCTGCCCTTTTACATCAAAGAACTACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTC  
TATCTATCTGGCTGGCAGGGAAGGAAAGAACTTGCATGTTGGTGAAGGAAGAAGTGGGGTGGAA  
GAAGTGGGGTGGGACGACAGTGAAATCTAGAGTAAAACCAAGCTGGCCCAAGGTGTCCTGCAGG  
CTGTAATGCAGTTTAAATCAGAGTGCCATTTTTTTTTTTGTTCAAATGATTTAATTATTGGAATGCACA  
ATTTTTTTAATATGCAAATAAAAAAGTTTAAAAACTTTTTTATAAAAAATTATTTNTTTTNTANTTT  
NTTTTTTANTTTNTTTNTTATTNTTANNTGGTTCTTNCCCGGGCGGCCNTTTTAAAGGGGGAATTTCA  
NCACANTTGGCGGCCTTNTNNTGGNTCCNAGCTTGGGACCAAACTNGGGGNAATNATGNGCAAA  
GCTNGTTCCCTGGNGGAAATGTTTTCCNTTCNAATTCCAAAAANATTCAANCCGGGANNTTTATG  
TTNNACC

SEQ ID NO: 752 CGAGGTACTGGATGTGGTTGCCCCCATTTGTGTGTGTGGTTGTGTGTGTGTGG  
TTGTGTGTGTGGTGGCCACAGCTGAGCCTCTGTACCAGAGAAGGCTGAGGCCCAATGGCACACC  
TCAGAAACCTACACCCCGAGGCTGGACGGCTGGACTCCTGAGCACAAGCTCCCTCTCGACCCCTT  
GCCAGACAGTTTGTCTCCAATTTCAAACCTGACCTAAGGCTCTTACTCCTGGATTTTTTGTTTTTAAA  
CCTTCTCCAGCCAGTCTTCGGGAGGGCATGATTAGAGAAGTGCTCCTTTGCTGATGGAGGAGGG

GACCTAAGGAAAGAAAGGTGGATCCAAGTGCCTTCTCTCTAATTGATCCCTCCCAACCAAGTTTCT  
 TTGCCTCTCTTTCTTNTACCAAGGCATGGTTTTTACTCTCTGCCCTTCTGGCTTCCAACATTTTAAAA  
 ACTNGTNAATGCACCCCATATGGACANTTTAAGTCCAAGCCCTNTNAATAGCATTCAAAANGTT  
 CTGGAATCCCAAGCACCCCTCTGAAAANCTNAGAAAATNGCCTTTCAACTTTTGGGTGAAGCNTG  
 GTTGAACTTTGGCTGTTCAAAGGTGCNNAATTNTCTTTGANGGTTAAACAANGCCANTTTACTN  
 TNNAAGACTNGGACCTNNAAGGGGGGGTCTTTTATTCCTNGACNCCAATGAACGATTTTCC

SEQ ID NO: 753 ACGCGGGGGTCTCATTGAACTCGCCTGCAGCTCTTGGGTTTTTGTGGCTTC  
 CTTCTGTTATTGGAGCCAGGCCTACACCCAGCAACCATGTCCAAGGGACCTGCAGTTGGTATTGAT  
 CTTGGCACCACCTACTCTTGTGTGGGTGTTTTCCAGCACGGAAAAAGTCGAGATAATTGCCAATGAT  
 CAGGGAAACCGAACCCTCCAAGCTATGTCGCCTTTACGGACACTGAACGGTTGATCGGTGATGC  
 CGCAAAGAATCAAGTTGCAATGAACCCACCAACACAGTTTTTGTATGCCAAACGTCTGATTGGAC  
 GCAGATTTGATGATGCTGCTTGTCCAGTCTGATAAAACATTGGCCCTTTATGGTGGTGAATGATGC  
 TGGCAGGCCCAAGGTCCAAGTTAGNATACANGGGANAGNCCAAAANGCTTCTATTCAAGAAGGA  
 GNGGCTTCTATGGTTTTGACAAAAATGAANGGAATTTGCNAAACCTNCTTTGGGAAGACTGGTA  
 CNNATGCTTNGGTNACAGGNCCGCTTNNTTTAATGACTTTAACNTTNGGGTTCCAAAN

SEQ ID NO: 754 GGTACCGACCATAGAGCAAGAATCAAGATTCTGCTAACTCCTGCACAGCCCC  
 GTCCTCTTCTTTCTGCTAGCCTGGCTAAATCTGCTCATTATTTAGAGGGGAAACCTAGCAAACT  
 AAGAGTGATAAGGGCCCTACTACACTGGCTTTTTTAGGCTTAGAGACAGAACTTTAGCATTGGCC  
 CAGTAGTGGCTTCTAGCTCTAAATGTTTGGCCCGCCATCCCTTTCCACAGTATCCTTCTTCCCTCCT  
 CCCCTGTCTCTGGCTGTCTCGAGCAGTCTAGAAGAGTGCATCTCCAGCCTATGAAACAGCTGGGTC  
 TTTGGCCATAAGAAGTAAAGATTGAAGACAGAAGGAAGAACTCAGGAGTAAGCTTCTAGACCC  
 CTTGAGCTTCTACACCCTTCTGCCCTCTCTCCATTGNCCTGCACCCACCCAGCCACTCAACTTCTG  
 NTTGGTTTTNCTTTGGCCATANGAANGGTTACCAAGTAGAATCCTTGCTAGGGTGGATGGGGGNCAT  
 ACATTCCTTTAATAAACCATGGGGACTGGGAATATATAGGCATGAACCCTGAGCCTGGGCCTGAA  
 ACGGTTTTTTA

SEQ ID NO: 755 ACGCGGGGGACGAACACGTGACGCGGTGCGGGCGGACCACTGCAGACTGAGC  
 GGTGGACCGAATTGGGACCGCTGGCTTATAAGCGATCATGTTTCTCCAGTATTACCTCAACGAGCA  
 GGGAGATCGAGTCTATACGCTGAAGAAATTTGACCCGATGGGACAACAGACCTGCTCAGCCCATC  
 CTGCTCGGTTCTCCCCAGATGACAAATACTCTCGACACCGAATCACCATCAAGAAACGCTTCAAG  
 GTGCTCATGACCCAGCAACCGCGCCCTGTCTCTGAGGGTCCCTTAACTGATGTCTTTTCTGCCA  
 CCTGTTACCCCTCGGAGACTCCGTAACCAAACCTCTCGGACTGTGAGCCCTGATGCCTTTTTGCCA  
 GCCATACTCTTTGGCATCCAGTCTCTCGTGCGGATTGATTATGCTTGTGTGAGGCAATCATGGTGG  
 CATCACCCATAAANGGAACACATTTGACTTTTTTTTTCTCATATTTAAATTACTACCAGAATATTTA  
 AAGATAAATGATTTCGNNNNNANNNNNNNNNNNNNNNNNNNNGGTACCTTGGGCGGGAACACNCT  
 ANGGGA

SEQ ID NO: 756 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTCTTTCAAAGATTTTACTAAATCAT  
 TTTTAAACAAAATATACATTAAATCTCAGATTTACAGAATATAGAAATAATTTATCCAAAGAAAT  
 TTGCATTTAAATTTGGTAATATTGCACCCAACAGTATGTCTTTGACACATTTGCATTGNCAATCTTT  
 CCCACAATTTGCAAAAACAGGAGAGAAATCTGAAACTAACAGAATTACACAAGCTAAGTTTTCTG  
 TAAAAAAGAAAAAAGTTTACAAATTTTTATTTACAAGTTAAGGAAAAGTTGTAAACGTTTCAAGAT  
 TTTACTTCCACAGAATAAAAAAGCCNTACATTCTTTTATCATACCTAGAAAAATGAACATNTACNCAT  
 TGNGATTACCTCATAGGGAATTCACAGGACGTGATTTGNGAACATTACAGCCCAATGGTAAAA  
 AACAGAATTTCTCGAACTNTGGGAACCTAGNNGGAAAAAGACNCCGAATGAAAAAAGNCTGGAAA  
 AAAATCTTTTCTCTTGNAGNTTCNTAATACCTAAAGTTTNTTAAATTTAAAGGNCCNCCNTTCC  
 CANCTTAGGAAGGGGGTGCANAACCTTTTTTNGAATANACAGNGAAAAAAGGGGGNTTGTNAG  
 CCCCAAAAAAGTTGNTTTGGGGGCCAAANTNGTTCAT

SEQ ID NO: 757 GGTACTATAATGGTCCCATCTTAATTTGAAAGCGTTTGAGAATCTTTTAGGA  
 CAAGCACTGACGAAGGCACTCGAAGACTCCAGCTTCTGAAAAGAAAGTGGCAGGGACAGTGGCT  
 ACGGTGACATCTGGTGTCTGAACGTGGAGAATTTCTTGCTCCTCCAAGGCACCATAAGAGAGAA  
 GATTCCTTTGAAAGCTTGGACTCTTTGGGCTCGAGGTGATTGACAAGCTGCTCCTCTGATATCAG  
 TTGAGAGGGGGCGTGAAGTTTTGAAAGTGACACAGATTCGGAATTTACATTCAAGATGCAGGA  
 TTATAATAAAGATGATATGTCGTATCGAAGGATTTGGCTGTTGAGCCAAAGACTGCGTTACCCTT  
 CAATCGTTTTTTACCCAACAAAAGTAGACAGCCATCCTATGTACCTGCCCNGGCGGGCCG

SEQ ID NO: 758 GGTACTATAATGGTCCCATCTTAATTTGAAAGCGTTTGAGAATCTTTTAGGA  
 CAAGCACTGACGAAGGCACTCGAAGACTCCAGCTTCTGAAAAGAAAGTGGCAGGGACAGTGGCT  
 ACGGTGACATCTGGTGTCTGAACGTGGAGAATTTCTTGCTCCTCCAAGGCACCATAAGAGAGAA



GATTCCTTTGAAAGCTTGGACTCTTTGGGCTCGAGGTCATTGACAAGCTGCTCCTCTGATATCACG  
TTGAGAGGGGGGCGTGAAGGTTTTGAAAGTGACACAGATTCGGAATTTACATTCAAGATGCAGGA  
TTATAATAAAGATGATATGTCGTATCGAAGGATTTCCGGCTGTTGAGCCAAAGACTGCGTTACCCTT  
CAATCGTTTTTACCACAACAAAGTAGACAGCCATCCTATGTACCTGCCNGGCNGGCCGTTAAGGC  
CAATCCANCCACTGGCGGCCGTTCTATGGTTCCNACCCTGGTACCAACTTGGGGNAAACATGGCAT  
AACTNGTTCGGGGGAAATNTTCCGNTACAATCCCACAACCTCCANNCCGAACCTTAAAGGN  
AAACCNGGGGGCCCAAGAGGGGGCNCAACTNCATTNN

SEQ ID NO: 759 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTAGNTATTCTTGCTTAATCATTTTTAAA  
ATTATTTCTGGATTTACCTACTCACCATATGAGCTGGGGTTTCTGACCTCCCTCCTGCACATCACGC  
CCTCCTTGCTGTGTTGCTTTTTCTAGCTCTGTTTCCCTGCATCTTGGGTTGAGGTTAGCAGACTGGCT  
ACCCCTTTCCAGAGCTCTGTTCCCTTTCTTTGCCACTGAAAGGAGGTTTGCTGTCTTACCGTAG  
CGGACACCGTTTGTATTAAATAATACCTCTGGACACATCTCCTTGTGTAAATCAATAAGCAAAT  
GCAATACAGGGAAAGACATACGGATTGGAAGTTGT

SEQ ID NO: 760 GGTACTTTTAATGGTGGGAATTTACAGTAGAAGCATCCTTTGCTGAGTTATAC  
ATTCCTTTATCAATCTCTTTTGATACAACATTTAAACAAGTAGCTTCAAGAAACCACTGGTGT  
GAGGATAGTATTTCTAAATAGCATTCAAGAACAGAGTATTATTGCACAGATCTGAAGATCAAAAA  
AAAGCTCAAGGAAATACAGATCGGAAGTGCTGATGAGTTATATTATTGAAAACCAACTTTTAA  
GGAAGTGCTAAGATCAGTCACCCATGTGAATAAGAAGCCAGGAAAGGAAAGATGGGGAAGCCCA  
GATCACCAGGCTTCTATTAAGGAGGAAAGCAACAGAGGAAACAGTGAAGGGGAACAGGAGGGG  
TAGCAAAAGTGTTACAGAAAAACCGACTGGATAGACAAAACCTGCAGAAAGGGTTGTGGGGGA  
GAACTGAAAGGAAACCAATCCTGACATGTCTTAAGTNAAGAAGGNNGTTAAGAAAACAATTTT  
TTATNGCCTTGACATTTCAANNCTTTAANTCCCTTTTGANANCCCCAAAAAATTTTGTNTTAAGG  
AAAATTTTTTTGCCCTTCCGAAAAAAAATTTT

SEQ ID NO: 761 GGTACTGGGATTATATAGGCATGAGCCACTGAGCCTGGCCCAGAAGCGTTTT  
TCTCAAAGGCCCTCAGTGAGATAAATTAGATTTGGCATCTCCTGTCTGGGCCAGGGATCTCTCTA  
CAAGAGCCCCTGCCCTCTGTTGGAGGCACAGTTTGAATAAGGAGGAGGAGGGAGAGAGAA  
AATGTAAAGGAGGGAGATCTTCCCAGGCCGACCATTTCTGTCACTCACATGGACCCAAGATAA  
AAGAAATGGCCAAACCCCTACAAACCCCTGATGTTTGAAGAGTTCCAAGTTGAAGGGAAAAACAAAAA  
GTGTTTGATGGTGCTAGAAAGGGGCTGCTCTCCAGAAAGCTAAAATTTAATTTCTTTTTCTCTGA  
GTTCTGTACCTGCCCGCNGGCGCTCG

SEQ ID NO: 762 ACTTTTTTGTTTTATTCTTTCTCTAGCTTATCCCTGCACAATTATTAGAGTGAA  
TGAAAAACCACTTTCCTGCTTTCCATTGTTATAAATTCTAAGCTTAAGATAAAAGTGGTTCTTTACA  
TGACTGAATCAATTACAATTTATGGGCTAGAGCCAAATAGGTTGAAGACAATCATCCANACAGAT  
CAATGGAATAGGAATTTCAATTGGAAATGTANAACACTTTCCCAANAATGGCATGACTTTCTCTGTT  
TTTGAGAAGAGTTTCATNTGCTGGACCACATTTAGCTTTNANTGTTGTTTTTTTCCCATNGTNCC  
NAAATTTTTANCTACANGNGGCCNCACTTTTACGTNGCCTACAACCTGTAGGTTTTGCAAAAGAA  
NNGTAATNTTCTTTGGTNTTCGCTTAAAGATATCTCCNCATTCGATGGGTACANTTTNATTTTGG  
ACTATAGGGGGACCTCTTANTACCTTTTTTGGTTATAANTAGNGTTTATNAANTTCTINTGNCCAAT  
TNTCNTNAANTGTNATTTNTNNNTTATTGANNNTTANAGNTTCNTTCTNATATTTACTTNNTAG  
TGCNTGAANATNTTACCNCITGNGNNCNTCCCCCTCTNGGNTNAANTCTNTTNTCTCTNG

SEQ ID NO: 763 GGTACCATAGTCCCAGCACTTGGCCCAGGGTCCTAGACTGCTGGGTAGGTCC  
TCAGAGGTATCTGAAGTCATGTCTACATTGATACAAAGAAACGTATGGCTCATCTGTAACAATACG  
TAGAGAAACACAATTAAGAACAATAAGGATCACAGAACTCGAATGGATCATTATAACTTCAAGAC  
TCAGTTTAAGTTACCAAGAAATGCACTGTAGCCCTTTTCTCACTTTTATATTAGATCAGGGCTAAC  
CCTAACATGTTCTTAAATGACCACTTATGGTCATTTAAGTGGTCAATGGGAATATTGCAGCAGAGA  
ACTGCTGGTTTGTAAACGT

SEQ ID NO: 764 ACAATAGATGCAACGCCAAAATGAGATGAAAGAGAATTCAGAATAAAATTC  
CGTCCCTTGAGCATATTCCATTGGCTTTCTCTGCTCCTGCTTCCGCTGGTATCTGGCCTTCTGGTA  
GAAAATGATGATGGTTACAGTGACTATATTTAGAAGAATGACGAGCAGTATGAGCCAGAACGAGT  
ATCCGTAACCTGTGGGTGCTTCCCTTACTGGTGGTTGCCGGGTAAAGCATTTGGAACAACCTCTCGG  
AGAGTTGGTTGACTGCGTGTTCGCCACAAACAGTATCATGGTCACAAAAACGAAGGATGCTGCA  
AACTTTTTCTTTGGTTCTGCAAGTCCGTGACTCAATTCTTCACTACTCTCCCCACGAAAAAGTCCCG  
TAGTGATGAAAATGCTCCCATTTGAAGCCANANNNTTTAANAGAATTGNCCTCGGCCGCGACCCC  
CNNT



SEQ ID NO: 765      ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTINACTNGTTTGTGGATGCTCTTTCCA  
GCAGACTCATGTCAGGAGGCTAGGAAGGGATTCTTGGGGACCACTGGATGCTGGTAGTTAAATG  
CCAGGAGGCTGAATGGACCTGAAGATGGAGGAGACTCTGCAGTCTTGGTTCAGCCACCCTTGGGT  
GCTTGCCACCCTGCACTNTAGCAGGATTGATGGTCTCTGGATTTGTAGCTGTGACCGGTGATGGT  
GAATGCTCGGTGGTTTGCAGTGGAGAGGCCACATGGTGGCNACTGAGGCCCTNTGGGGTGAGGT  
TGGCTCATGATAGCTNTGAAAAGTTGATGGCACAATTGAAACAAGGACGTGNAAGTTTGGAACTT  
TNCCAGGGTCTTNCANCCCCGANNTAAACNTCCCAACAATTGCANCTTACTTGTNTNAAAGCTT  
TTAGGTTTTTTNNAGTTGGGCTTAGGGACTNGNGGGCAAANAAAACTTGTNTTCTTCNTTANN  
AANCTTTGNNAGGAACCTTTTTNTTTTTTGGNGNAAAATTNGGAAACCCATAAACCCNGTTTC  
CTANCCNGCTTGNAAGGNATTAACCTTTTNTNGNAANGNAATTTTCCN

SEQ ID NO: 766      CGAGGTACTGTTACCCGGACATACCTATTCCTTCACAAGTTCTGGTTCTTCAG  
TGCTGCTTACTATTTTGGTAACTGGGCCTTCTTGGGGTATTTTTGATTGGATTAATTGTATCCTGTT  
GTAAAGGGGAAGAAATCGGTTATTGAAGGAGTAGATGAAGATTCAGACATAAGTGATGATGAGC  
CCTCTGTCTATTCTGCTTGACAGCCTTCTGTCTTAAAGGTTTTATAATGCTGACTGAATATCTGNTA  
TACATTTTTAAAGTATTAACTAACATTAGGATTTGCTAACTAGCTTTCATCAAAAAATGGGAGCAT  
GGCTATAAGACAATAATTTTATTATATGTTTCTGAAGTAACCATTGNATCATAGATTAACATTTT  
NAATACCATAATTATGCTATGTTAATATAAGACTACCTGCCTTTGTGAGGGAATGTTTGTGCAAAA  
TTTTNCTNTTATNGNTAATAGNGGTTAAATTGAATAAAAAATCTCCAGAATTAATNTNCCNTTT  
GNCACTTTTGGAAACATAATAAATTTTTGGATTNNNGCAAAAAAAAAAAAAAAAAAATTCCTN  
CNCNGGCGGCCTTCTNAAGGNCAATCCCCCTNGCGGCCGTTTTTNGNTCCNNCCGGCCAACTNG  
TNANTNAGGGCTAANTTTCCNNNNAAAANG

SEQ ID NO: 767      ACTCTATAAATCTAGTGGAACATTTCTGCACAACTAGATTCTGGACACCA  
GTGTGCGGAAATGCTTCTGCTACATTTTTAGGGTTTGTCTACATTTTTTGGGCTCTGGATAAGGAAT  
TAAAGGAGTGCAGCAATAACTGCAGTGTCTAAAAGTTTGTGCTTATTTTCTGTAAATTTGAATAT  
TGCATATTGAAATTTTTGTTTATGATCTATGAATGTTTTTCTTAAAATTTACAAAGCTTTGTAAATT  
AGATTTTCTTTAATAAAATGCCATTTGTGCAAGATTTCTCAAAGATTAGGTATATATTTAAATGGA  
AGAGAAAATATTTTTATGGGAGAAAAATCATTTGAACCATGAAATTTTCATCTTTTNAATAACATCC  
AGTACTTTTTTTTTTTTTTTTTTTTTTGGATATAAGGGGGAAGGANANTGNCTTACTTCTTAAGTT  
NCATGGNGGATATNCAAGTTGTCCCCGCACNTTAGTTAAAAAACTTTCTTTGCCCTNAATTGTCT  
TGGGNCCCTGNCCAAAATCAATTTCCCTCCATGGGAGAAATTTTTTNGAANTNGCANCCCTTCCCT  
NAATCNGGTATTCCACCTTTAANAACAATTGGNTNGNTCNGCCTTTGTAAAAATTTGTAATTAGGA  
AATTAAGTNTCNACTTTTTTTTNGTCAAAAAG

SEQ ID NO: 768      GGTACGCGGGGATTTGTGGTGAGATTCTCTCCCAGGCCACANGACATTTCTCG  
CTCGGAACCTTGTTTACTAATTTCCACTGCTTTAAGGCCCTGCACTGAAAATGCAAGCTCAGGCG  
CCGGTGGTCTGTTGTGACCAACCTGGAGTCGGTCCCGTCCCGTCCCCAGAACTCCAACCTGGCA  
NACAGGCATGTGTGACTGTTTCANCGACTGCGGAGTNTGCTCTGTGGCACATTTTGTTCCTCCNGT  
GCCTTGGGTGTCAAGTTGCANCTGATATGAATGAATGCTGTCTGTGTGGAACAAGCCGTTGCAATG  
NGGACTCTCTACANGANCCGATATGGCANTCCTGGAACATTTTGTGATGACTATATNGCAACTNTT  
TGCTNTTCTNAATTNNACCTNCCCGGGCNGCCCGTTTAGGGCGAATTTNNNNACACTNGNCGGNC  
GTTANTGTGGGATCCTAGCTCGGTTCCAAGCTTGNCNTTNTCAATGGTCCTANCTNTTTCTCTGTG  
GTGTAATNGTATNCGNTNNNAATTCNCCNATNTTTTTATCCCGNANNNTATAANTGGTNAANTC  
TTGGGGTGNCTTAANGNGGGANGCTTCCTTANATTATTCGNTNGGGNAACGTNTCTCCNTTNAAT  
GTNGAAANCTTNTGTCAATCTTTTATATGANACGNCNNACGCNNTGGGG

SEQ ID NO: 769      CGAGGTACTTTTAAAGAAAAGTCCAATGTTACAAAATCAAATGCTTATATTCA  
GACTGGCACACTTTTTAAATAAAAACTCCATACACCTCAGACATATAGCACACATGGAGACAACT  
TACTAATTGTGTGTAAGTATGATACAATGAATGAGACTGCCTGAAGTCTAGTAATCAAAGCATGCC  
ATAAGGTGAATGATTGTGGTTAAACACAGCAAAATAATTGTCACAAAACCTTTCAAGGCCTAACAA  
ATTAGAATTTTCCAATAAAAAATATATTTTTTTCAGATGTTAATAAGACATATCAGTAGAGACAA  
AATTAGGATTTTGAAGTAATGCAATAAAAAGATGTTGGAGGCAAAAAAAAAAAAAAAAAAAGT

SEQ ID NO: 770      GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCGGGTNGNCTGATTTTTA  
TTATTTAAAAAATGGAAAAACAAAAGNGCATTTTTCATTCAATAAATGTTCCATCCTTATTTAGT  
TTTGTGCGGAAAGNGAAGTCCATGACTTTANAATGATAGCAATTTATCAACCAAAGAATCCGTNT  
TCACACCGTTTCAATAACTGCAGCAATTTCTTGAAGTGTCTGTAAAAATTTNGAACTGTGGAAT  
CGTCATTTCAAAGCACTTGGTCTTTACTTGGCCTGAATGATCTGCCACTTTTAGCATCACTGCAACG  
TAAGGATACTTAAAANATCTGCAAGTGTNTGAGCTCACAGCCATACCCAGTTTCCAAGTGAATCT  
ACAAGCTGGTTGGNGACATCGGACTTAGCATCCAACGGNGGNTCGCTGGACNCCTCCATGGGCAG  
CGTNTGGGACTTGGCGCCGGACTNNAACCCCCCGTCTGCCCGGCGGGCGGTCAAAAGNGAAA

TCCACCNACTGGNGGGCGTTCTATGGAANCNAACTCGGGNCAAACCTGGNGGAANNATGGCAAA  
CTNGTTCCTGGGGAAAAAGGTNTCCCTCCAATTCNCAAATACAANCCGACCTAAAGNNAANCNG  
GGGGCCCANAGGGACAACCCAA

SEQ ID NO: 771 GGTACCGCTGAAGACACCCAGAATGAAGGAAAAAAGACAAAAAAGAATAAA  
ACAGCTTTTAGTAACGTTGGAAGAAAAATTAGTCAGCGAGTTATTCACCTATTTTGATGAGAAGGGC  
AATGATTTGGGAAACATGCACCGAGCAAATGTGATTAGACTTATGGATGAGCGAGACCTGCGACT  
GGTTCAAAGGAACACCAGCACAGAACCTGCAGAGTATCAGCTCATGACAGGATTGCAGATCCTCC  
AGGAGCGGCAGAGGCTGAGGGAGATGGAGAAGGCGAACCCCAAACTGGACCAACCCTGAGAA  
AGGAAGTATTTTGTCTTCAAATATTGGACAACATGATTTGGACACAAAGACTAAACAGATTTCAG  
CAGTGGATTAAGAAAAAACCTAGTCCAGATTACCATAAAGAAAGGGAAAAATGTAGACCTGTC  
ANAAAATGAAATGGAGGAGATTTTCATCAATACTCCAACCTATCCTGGAATACTCENNTCTCATNTAG  
GCCCAGCTTGTGANGAGGAAACCTTNAATGGGGGTCNCTCGGCCTTGCCCAAAATGGGGGAAGGCTT  
TAAGAACTCAGAAAACCCGGAAGAACCCTTTACCAANNCTGGAATGTTAGGGATNAAGTTTCCTC  
CNAATTTATAAAAAACCTCTTTTAAAA

SEQ ID NO: 772 ACTGTATATCCATATGGCACATTTATGACTTTGTAATATGTAATTCATAATAC  
AGGTTTAGGTGTGTGGTATGGAGCTAGGAAAACCAAAGTAGTAGGATATTATAGAAAAGATCTGA  
TGTTAAGTATAAAGTCATATGCCTGATTTCTCAAACCTTTTGTTTTCTCATGTCTTCTGTCTTTA  
TATTTTATCACAACCAAGATCTAACAGGGTTCTTTCTAGAGGATTATTAGATAAGTAACACTTG  
ATCATTAAAGCAGGATCATGCCACTCATTTCATGGTTGTTCTATGTTCCATGAACCTAATAGCCCC  
ACTTATACATGCGACTCCAAGGGGATGCTTCAGCCAGAAAGTAAAGGGCTGAAAAAGTAGAACA  
ATACAAAAAGCCCTCGTGTGGTGGGAACTGTGGCCTCACTCTTATTTGCCTTCATTCAAAACAGTTG  
GCCCTTINCATGACGAGGATCTCTACAGGTAGGGTAAAATCTTTTCTGTGCTATCAGCCAGAAATA  
GTTTTGGGCTTGATATGATTTNAAACANATTTGGCTGGCACCANGCAAAACATACAAATGCTGG  
CTNTTCCAAACCCTTAGAATCTCCNCTTTTTTAAATANTTTCCGACCTGGGCACNCCCTANGGGA  
ATCCANCTGGGGCG

SEQ ID NO: 773 GGTACGCGGGATCTGGGAGAATATTTAATGGAAAATCGCTTGTTAAAAACCT  
GACACTTTTAAACAGTGAACAGCGTTCTGAGTGTGGACGAGTAGCCAGTGAAGATAATGAATGTCG  
AATGTGACTGACTAGCAGCTTCATTTTGAATGAGGGTCTGCTGTCTGCCATTGATAGAGGCCAGAT  
TGTCTTGGAAGTTCCAAAGTTGCAACGATTTCTGGCTAGTGCCACGAGGTTTACTTGACTGTTGTG  
TGAAAAGCTGATAAGAAAACCATCCAGAAAAAAGCTCTTCGTTTACAAACATGAAAATAAAACA  
TGTAATTTTCAAAAAAAAAAAAAAAAAANNAAAAAAAAAAGGT

SEQ ID NO: 774 GGTACTTTATGAATTTGGGGTAGGTAAAGTTTGTATTTTATCTTAAACATGTT  
TTCTATGATGAAAAGGAACAAAATTGTAAAAAATGAGGATCTTCCCTCTAAAGGTTTCAAAGCGT  
TAGAGGACATGCAATTAATGTTGTACACCTTGAACAATGAGCCTCTTGAGTTTGTAGGAAGGGC  
AGACCGCTCCATTACCAACAACCTTTGGGGTAGAAAGCACAGCTCTCCTCTTTTACCCAGCACAAA  
TGCAATCCTGATTATAAACTATTTGTGTTTCTAAATACAACCAAAGGAAATCTTAGAGAAACATA  
AATTAGAAACCTCTTTTATTAAGGGGAAACAACAAAAAAGGTGCTTTTTTAAAAAAAAAAAAAGAA  
AAGAAAAGAAAAAGAAAAACAAGCTGTAAAACCATGAAGTTAAAAGAGCTGGGTCTGGAGCAG  
ATGTGCATAATAACTAGTTAAGTGCTCCCAGCANGATCAGAAACAGCTTTGGGGGACCAGANGAA  
TATGGGTTGGNGNTGTTCANAAAAGCNCCAGTTATACTCTTCCATNAAAATGATGGGCACAGTGTT  
C

SEQ ID NO: 775 GGTACCCAATGAGGAACCTAAAGTTGCAACAGCTTATAGACCCCCAGCTTTA  
AGAAATAAACCAATCACCAATTCCAAATTGCATGAAGAGGAACCACTCAGAATATGAAACCACA  
ATCAGGAAACGATAAGCCATTATCAAAAACAGCTCTTAAAAATCAAAGGAAGCATGAAGCTAAG  
AAAGCTGCAAAGCAGGAAGCAAGAAGTGACAAGAGTCCAGATTTGGCACCTACTCCTGCCCCACA  
GAGCACACCACGAAACACTGTCTCTCAGTCAATTTCTGGGGACCCTGAGATAGACAAAAAATCAA  
GAACCTAAAGAAGAAAAAGCNATCGNACAACCTGAAAGAACAAACAGCAACTGGAAAAACAGNTN  
GAANAAAANTCAGTTGGGAGAAAAATTNTGAAAGAAACAGCCCTTCTTCAGGANGCTGGAAGATTT  
GGGAATTGGGGTTTTTTAAAAAGATTACCGGAAAGCANGGTTGNTGACCAGAAATCAGTNGCAAA  
CNCATTCTTNTGGTAAACCCNTTNGNTNCCCAGAANNNTTCTTGNGCCCNCTTTAATGTCAATN  
TTTAATTTAAACC

SEQ ID NO: 776 GGTACGCGGGGAGGCTCGGACCGGCGCGGAGCTGCTGCAGTCCTTCGCGC  
CCTCCTCGCCCTCCCCACCGACATCATGCTCCAGTTCCCTGCTTGATTTACACTGGGCAACAGTGG  
TTGGAATGTATCTGGCTCAGAACTATGATATACCAAACCTGGCTAAAACCTGAAGAAATTAATAA  
GGACTTGATGCCAAGAAGAAACCCCTAGTGCATGAGACTGCCTCCAGCACTGCCTTCAGGATA  
TACTGATTCTACTGCTCTTGAGGGCCTCGTNTACTATCTGAACCAAAAGCTTTTGTTTTCGTCTCCA

GCCTCAGCACTTCTCTTCTTTGCTAGACCCTGTGNTTTTTGCTTTAAAGCAAGCAAAATGGGGCCCC  
AATTTGAGAACCTACCCGACATTTTCCAACATACTCACCTCTTTCCATAANCCCTTTCCACTGCATG  
GGAGGTTTAAACTGGAANTATGGTGCTAGATTATTAACNATGACTTTTANTGANATTNNAATT  
ATTACATCNNNCCNNAACNTNNNAACANTNGNTGCNNANANTTTATTTAGAAAAANNATGGG

SEQ ID NO: 777 GGTACCTTGGCTGGCATTATTGGAATGAGGTTCTACCACTCTGGAAAATTCAT  
GCCTGCAGGTTTAATTGCAGGTGCCAGTTTGTGATGGTCGCCAAAGTTGGAGTTAGTATGTTCAA  
CAGACCCCATTAGCAGAAGTCATGTTCCAGCTTAGACTGATGAAGAATTAAAAATCTGCATCTTCC  
ACTATTTTCAATATATTAAGAGAAAATAAGTGCAGCATTTTTGCATCTGACATTTTACCTAAAAAAA  
AAGACACCAAACTTGGCAGAGAGGTGGAAAATCAGTCATGATTACAAACCTACAGAGGTGGCGA  
GTATGTAACACAAGAGCTTAATAAGACCCTCATAGAGCTTGATTCTTGATATTGATGTTGNCCTT  
TCTTTCTGCATCTGTAGGTAAATCTCAAGGTAAAAATGGTANGGGTCAACTTTNAAGGCTCTGAAA  
CCCCATTTCCCTGCTCTGAGGAACAGTGGGAAAAAAGTCTTTTAAGAGATTTACNATATCNGNNC  
NTTGTCTCTTAAANCCAANACTGCTTTGGAATTTTTTTAAGGGAATCAATGAAATAAGGTTT

SEQ ID NO: 778 CGAGGTACAGGATAATATACTCAGATATTTTTAAAAATAAACTACTTAATAAT  
AAGAAATTAGCCATACCACATTGNTCCATTTGCTACAAGAACAATTTGGCAATGAAGACTATTTA  
AAAGAAATGCTCANCTCTACAGAGGGTGGTGGCAGGCAACACTTTTCCATTACAGAATAACCTCT  
ATTCTTCCATGATACATATTCCTGTGGAAAACTTGTGAGGGCCAGGGATGAAAAATANAGCTTG  
NCCTAATTAGCTAACTGTAGGTTCACTTAACATCTTTGGGAAGGACCCAAAAATCTGGCCATTAT  
TTTCTAAACATCTGCAAGCTGCAAAAATTCCTTAGTCCCTCAGCTATAGTTTCTGTANATATCTTA  
NAGCTGGGACAGCTNCACTGTGACTCCTCCTCAGCTATGGGGTGGGTGCTAGTCATCANAGTCTGG  
GAATGTCTCATAAGAAGTAACCACGGGAGCCTTTGGATGCATAANAAGCCGATGCCCCAGGTGGG  
AAAATCCGGGTAGGAACCCAGAAATGCCAATGNTCAGCACNGGGAACGGCCCGGCTGGCCCC

SEQ ID NO: 779 GGTACCAACATTTATTAATCCTCACAACACCCCTGTGAGGTAGGTGAGTATGT  
CCTTTAGAGTCGAGAACTGAGGCAGAGGTCAAGCAAACCTGCCCTGGGCCACAGAGCAGCAGAT  
GAAGGGCCTAGACCTGGATCCAGAAGCTAGGGCTCTCGGTCCAGCATTCCACTGGTGGACAT  
CACATGGGCTTATTTTACCAGCGAAGGTTACGTGAAGGACAAAAACGCATCAGCCAGCAACGGA  
AATCAACAGTTTCAAACAGCACTGGGGAACATGTCAGTTAAAGAGACGAAACGCTGACCAGCTCA  
TGAATGAGGCAAGACAACATGCGGCTGAGGAAGTGTGGAATCATCACGACTGGGGATTAGACCA  
AGAACGGGCGCTCAAGAGGGTTCAAGAAAATGTAAACAACTAGGAACCCATCCAAGGGGTGAC  
AGGCCCAAATGCCTACGGCTCCAAATGGTAGAAAAATTAGAAAAAGATAGNGGAAAAATNCCAC  
CCCCANCCTTACNTTTAACTAGGAATGCCATAANGGGNTAAGGTCAATCCCCCTTAAAAAG

SEQ ID NO: 780 ACGCGGGGGCTCAGAGCTCGGGGGCGGCGCTCAGAAAACATCTGGAGAAAA  
TGACCCATTGGTTTCATAGGAACCCATTAAAAAGCCACAGCTCCTGTGTCTTTAATTACTATGGTGT  
AGTCACTGGCCCTCTGCTTCAAAAATATGCAATGACTTGAGGTGATCCAGGGCAGACTCCTTGA  
ACTGTTCACTGATTTGAGCTGTAATCCAGAAATGATGAAGAATGCAGCAGATTCATATTTCTCACT  
TTTACAAGGTTTCATAAATCTTTGGATGAATCTACCCAAGAAAGCAAGTTACGATATATTCAAAA  
TTTCAAGTGGACTGATACATTGCAAGGACAGGTTCCAAGTGCCAGCAGGATGCTGTTTTGAATT  
AATTCCATGGGATTTAATGTAGCTTTATGGTATACCAAATATGCTTCAAGACTGGCTGGAAAAAG  
AAATATAACAGAAGATGAAGCAAAGAAGTTCATCGAAGCCTAAAGAATGNAGCTGGGATTACAG  
GCNCCACCACTACACCCAGATAAATTTTGNATTTTAATAGAGACGNG

SEQ ID NO: 781 ACACTTTAATATCTTTCCACCAAAGGCGCGACAGCACTCTGCCAGATCTTGA  
CTGGCCTCTGGTGGGCCATGGACGATGATCAACTGTCGTGGTTTCATCTGATTAATGATTTTTTTAA  
TGGAATCCCCATCAGAGCGTCCTTCATAGTCTATGTAGGTAAACCGGGCTTTTATTTCAATAGACT  
CTGTTGTAGAAATACATTTAGTAGGAACATCAGATAAATCCTGATCCATAGGTTTCATCTCCATTTG  
TCAAACAGATTCTAATTTGCTTTTTCTTCTCAGTAGCTTGAAGCTCTGGCACTAAGAAATCCTC  
TGGTTTGATAATCTCTCCATATTCATCCCATTTAATTCTTTCTTCTGGGGCAGGAAACATAGGATAG  
GACTTTTTTGCCTGTTTGAaaaaactTCCTTTACGACTGCCTTCACCTTTCATCATCAAGTCATGCT  
TCGCTTATGAGCTGATGGCTGGNCAATATCTTCTCAATATCACTCTCATCACTGGNATCTATATCT  
GGCTCTTTTGACTIONTCAAGCTTTTTTGCAAGCTTNTTTT

SEQ ID NO: 782 GGTACTTTAAGAAAAATTTAAGTGACAAATGTTAGAAGATGATGGATTGAAA  
AATATAAGTAATTTGGCATATTGACCCTATATAAAAAAGTGGGCATTTTGTAAACATTCCTTCAGGA  
AGAATAGAATTGTATGCTTTTTCTGCATGTTTGTGATCACTGTGAGTCTCAGATGATTATTCCAGT  
TTTCAATGATTTTTCAAATAAAAAGTCAATCAGCATTGTTATTTATCATTAAGGTATTGAACACTG  
GATTGCATGGTTGAATGTGCTTTAATCCACTACAAAATGTGCTTGTTATTGATAGGATCATGTTGT  
TAAATTGTAAATTTTCAAATAAATTGATGTTTTAAAAATGTTGAGACCAAAGTAGTATAGAAGTA  
TGGACTTAATTTAATTGTAAAATTATTAGAGGTATTTGTTGTAGAGCTTTATATATTAAAGAGAGG

TATTGGTGCATATAAAAAACAGTAATATTATTGGTATGCTTGGATTCTTGCATTTTCAGGGTAAGTAA  
ACCCTGAAGAAATCTTATTTAAGTATAGCTACNGCTGCCGTAGCTT

SEQ ID NO: 783 GGTACCAGGAAGTAGGACAAGTAATTTCAAAAATATAAAGGTGTTTGCTACT  
CAGATGAGGCCGCCCCTGACCTTCTGGCCAGAGAGACATTGCTGCCAGCCAGCTCTGCCTTCCC  
CATCTCCTTTCAGGACCGTCCACACCTTTTACTTGCTCAGTGCTGTCTGAAGATGCAGTTGCTGTT  
TGCAAACAACAGGAACACCAGTTAACTAATTAGGAAAAGAGGGAGATTTCCAGGCCTGGGTAA  
CTATATACTGTGACCATTGGAGGTAGAGACAGGTCTCAACAGTTGGAACCAGGAACTCTGCTGTC  
AGGTTGTGAGTTTGTCTTCTTCCAGCTTTTCACTGTGTGGGGGTCTTTTCTCTTATGTCAGCTCTT  
TCTATCACATGGCAGCTGACCTCTCACGCTCCACTCTACAGCTTGGACACCCAGTAGACCCTGAAT  
TTCCTCTCTCTAAAAGGTTCTGAGGGCTCATCTGGGCCAGGGGCCCTCTGTGCACTGTTAGCT  
ATGGCCACGGAAGCCTNCAAGCTGNCTGGAACCTCAGGTTGACCTGTT

SEQ ID NO: 784 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAATAAGTTATACTTTTATTTT  
TCANAAACAAAAATGACAAGNGGCAACTTGCCCTTGTAAAAGATTAAAGAGTATCAGAGTAATAA  
GCTATCTCTCATAGAGACAGAAGCCCCGAGAAACACTATTATACAGCTTTCAAAGGAAGTCCTATT  
AAATAAATCTCACAATAATATACTTTAACTCAACAAAAACAAAAACCACAAAAAATTATACAGAAC  
TCATAAAATGGTATATGAATGAAAAAGAGAAATCTTTAAATCCATACACTAGAAGTTCTCTATTAA  
AATCAAGGAGGCTCAAACTCTGTTTCTTTAAAAAATATTTATGGATATAATTGCGTTTTTCTAAATG  
TTTCATTAGTAATACTCAGTCTTGTAAGTCTGGTTCCATAGTAATCTATAATGTAACACTCTCCATC  
TGAAGGTCCAACAATCTGCATTGAAATAAGAATGAAATCAATTAGGCTCCCAATTCCACAAAAACC  
CTCAGGGCAAACCTTAAACAAACCCAAAGCAGGGTATCCANGNAAA

SEQ ID NO: 785 GGTACACTTAAGTTGAAGACACAACACTTGATCTGAAACAAGAAGTTTGTGC  
CTACTCAACAGCTTTGAAAGAGCACTTCCCAACGCTGCTAGTAGTCTTTGTTTTCTTCAGTGCTGTA  
CAGTGGTGTAACATAGCTCACTGTAGTCTTGAATTCCTAGACTCAAGCAATCCTCCACCTCAGC  
CTCCTGAATAGCTATGACTACAAATGTGCACCACCACACCCAGCTAATTAACAAAAAAT  
GTAGAGATGGGGTCTTGCTATGTTGACCAGGCTGGTCTCAATCTCTTGGCCTCAAACAGTCTTCCC  
ACCTTGGCCTTTCAAAGTGCTGGGATGACAAGCATGAGCCTCTGTGCTCAGCTCTTCTAATAGTTT  
TATAGTTTCAGGNCCTTACATTTAAGTCTGTTTCATCCATATTGAGTTCATTTTGTATGTGGNGACAGA  
CAGAAGNCAAGTTTCTTCTTATGTGTATGAAAATCCAGTTTTCCTAGCACATTTATTGAAGANGC  
TGGCTTTCCCATAAATGTTCTGGGTCTTTTGTCAAANANNANTGGNTGNAAAACCAAATTGGTTC  
TGACNCCTATCGGTTCTGAGGCTAGGGGNGGGCTTAAGNCCACCCGCCNTTTGATTATACTTGGA  
GGTTTTTGAAGCCGNGGGGGGCCCTACTT

SEQ ID NO: 786 GGTACTTNTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAATCATTTTATCTTTAATAATCTT  
GTAGTAATTTAAATATGAGAAAAAAGTCAAATGTGTTCCCTTATGGGTGATGCCACCATGAT  
TGCCTCACACAAGCATAATCAATCGCCACGAGAGACTGGATGCCAAAGAGTATGGCTGGCAAAA  
GGCATCAGGGCTCACAGTCCGAAGAGTTTGTTTACGAGTCTCCGAGGGGTAACAGGTGGCANAA  
AAGACATCANTTTAAGGGACCCCTCATAGGACAGGGCGCGGTTGCTGGGTGATGAGCACCTTGAAG  
CGTTTCTTGATGGTGATTTCGGTGTGAGAGTATTTGTCTCTGGGGAGAACCGAGCAGGATGGGCT  
GAGCAGGTCTGTTGTCCCATCNGGTCAAATTTCTTCAGCGTATAAACTCGATCTCCCTGCTCGTTG  
AGGNAATACTGGAGAAACATGATCGCTTATAAGCCACCGNCCNATTNGNCCACCGTTAATCTG  
NAGNGGTCCCCCGNGTCTGCCCCGGCGGCCGTNTAAAGGCGAATTCANCCCCCTTGCGGCCGNANT  
NATGGANCCACCTNGTCCCANCTTGGNGNANTATGGCANACTNTCCGGGGNAAATGTTCCNTCC  
ANTCNCAANTNNAACCGGAACAAAANTAAACCTGGG

SEQ ID NO: 787 GGTACAGTTCCAAAAATAATTAATTTTTTAAGGGAATTTTTCAAGACAAAAG  
GCAAATATTTTGTTCCAAACCTACAAGCAAATTCATTCCAAGAAGTGAATATAGTAGTAAAACATA  
AAGACTATTGCAATAAAAAAGTTTCCAGGCCAGTGACTTGTTTAATTACATTACATTTTACATT  
ACTTTGCCCTTTCAAAGCAGGAGTATTTTATAAACTGAAATAAACTTTAAATTGTCAAGCTGGTG  
CTTAAAGTTATGTAAAATCTGATTTAAGTTTATTTTCAATTTCCAAGGAAAGGACCCTGAAAGGAGGG  
GACAGGGAGCTTCTAGGTCTGTTGATAAACTAACTGATGCTTTGTCAAGTCTTCCAAGTTTTCAC  
TGTAATAATACTAACTAACAATGTAATTTTGCTGGTTTCTAATAGAGAATGAATGAATATATTAAT  
GNATTACATTGGTTCATGATAAAACAGNTTCTCTGATCTGGGAAATCTTCTCTTNTAGCCCT  
GCNAANACCCTTAAATAATATTTTAACTGGAGGCCGNGGCACTGGTGACCGTTTTGGCAATTTT  
GNANAAATGATANATTNCTGCCCGCGGCNCAAGGNGAATCACCCCTGGNGGCGTCTATGGACCAC  
CCGGNCCACTGNGGATATG

SEQ ID NO: 788 ACAGCTTTCTTCGTCTCCATGCTAAGAGATGTAAAAGCTTAAGGGTCAAAC  
AATACCAATTGTATAGGCTTCAAAAACCATCTAAGTTAGGGCATTCTCTAGTTTTAGCTAAGATAC  
ACCTGGAACACTGACAAGTCATCACTTACATAGAATAATGTGAAGTAAATTTTTGAAAAATAAA

TTTGTAGTGGAAACAATCCTGAAGGATAACACCAGAAGAATAGCAGGTTACCAGTAAGGTGTCAGCC  
AATTTGTTCCAGTCACTTTTGAATCCATGTTCTATAATCTAAAATTTATTCTCTTTCCCTAAGCTGA  
GAGCTTCCTATCATGTCTAGTATCTATGTTATGAAGAAAAGGAGACTTANGTGAGATGTTTTTATTT  
ATCGCAACTGCTGCATTAATTGCCTAGGACCTCAACAGCTTCATGAAAGTCTGGGAAATGTTTCATG  
CATAAGGGTATTGCCTTAGCTGACTTAAAATTGGCCCATCAATGGTCCTCGGNCGCGAACCCCC

SEQ ID NO: 789 ACGCGGGCAGCTAAGCTAAAGGAGAAATATGAAAAGGATATTGCTGCATATC  
GTGCCAAGGGCAAAAGTGAAGCAGGAAAGAAGGGCCCTGGCAGGCCAACAGGCTCAAAGAAGA  
AGAACGAACCANAAGATGAGGAGGAGGAGGAGGAAGAAGAAGATGAAGATGAGGAGGAAGAGG  
ATGAAGATGAAGAATAAATGGCTATCCTTTAATGATGCGTGTGGAATGTGTGTGTGTGCTCAGGC  
AATTATTTTGCTAAGAATGTGAATTCAGTGCAGCTCAATACTAGCTTCAGTATAAAAACTGTACC

SEQ ID NO: 790 ACGAAGCCATCGACAGCAGAGATGGAGCATCTTGTGCAGAGTTGGTGTCTTT  
TAAACATCCTCATGTTGCAAACCCACGACTTCAAATGGCCTCTCCAGAGGAGAAGTGTCAACAAG  
TCTTGGAACCCCTTATGATGAAATGTTTGAGCTCATTAAAGGTGCACTTATGCAGTGGGGAATC  
ATGACTTCATAGAGGCATACAAGTGCCAGACCGTGATAGTCCAATCATTCTTGCAGCATTCAGG  
CCCACAAAGAAGAAAAGTGGGCTCTGCCTGTCTATGATGACAGTAGCCGCTTGACCTTCGAGTGT  
GCCAATAATGCAGATCAACAGTTGGTAAAGAAAGGAAAAAGCAAAGTTGGGGACATGTTGGAAA  
AAGCAGCAGAGTTACTGATGAGCTGTTTCCGGGTCTGTGCCANCGACACCCGTGCTGGTATAGAG  
GACTCTAAGAAGTGGGGCATGCTTGTCTGGTGAACAGCTGGTTAAAATCTACTTCNAGAACA  
CAAACCTCATTATGTAAACCCCTAATTTGAACAATTTGACAGCTCAAACCTGAAAGACGNTTNA  
GCNCTGCNCAGAGAGTAAACATACAAATACTNCCTTGGACCCAGGNTTTTTTTGNAACCATTTNAAC  
AANCTGAGGATNCNTCGCCGCAAANGGGGAATCCAACNCTGNGGCGTAT

SEQ ID NO: 791 ACTAAAGGCTTTTGCATGAATTAGGAAGGAGAGTCTTGGGGCAGAAGCAATA  
GGGGACAACCTGTGCTGGTGTCTTTTGCAGGATGTGTTTACCAAAACATCTAATGCAACTATTT  
TCAGACTTTACAGTTTGTAGTGTTAACCTCTTTAGAAAAAGAGCAGCCATCCTTTTTTAGGAAGAA  
CTGTAATCATCCCCAGTTGATGAGGAGAAGCTCTTCTGTAGAGAAGAATGACAGCTGTGCTGGG  
CAAGCGATTGACATACTGTAGCGGACGCAAGTACGCGGGATAGCATACTTTGACATTTTAAACAT  
GATAGTCCATAACCATTTTGAATGCTGGGCAAACCTACATGAAGTTATTTATAATTAATTCACAGC  
TAATCAGGCATTTTGAAGCTTAATTGGATTCAAAAACCATAATGTTGGAATTTGGTAAAATTTTA  
ATGTTGATTTTACTGTGAAAAGGTTTTTATAAGATATACACACCCTAGTTTAAATGTTGTGCTTGG  
TGTGGATTTACAGATTTACTACAGGTATTCTGAACCAGGAACACAATCAGGTTTCAGGCCAGTTTG  
ATACTGGCTGCCCTAATTCTAATATGANAGTAGGACATCATACTAAATGTTATGTGCTGAGTGGGACTG

SEQ ID NO: 792 ACTTTTGTATTTTGATATGGACAGTTTATTCAATTGTCATACAGTTATTGACTTT  
TTCCAGCTGATTAAAGATAGTCAAGAAATTCTGCAATATAGCTGCCAAAATAGACAGCTACAT  
TTTTATGATATTGTCATCTTTTCTGTTTTTTTTTTCTTTTTTTCTTTAGCTATTTTACTTAAGCATAA  
TAGCCACAATAGGACATATAAAAGATTATAAATACAGAGCTTTATTATCCTGACGTCTTGGGTCTT  
TTAAGTATATACTTTTCTGAAAGGTATCCATTTTGTAGGCTTGGGTCTTCATGAGCATACGATTGT  
TTATTTTGTCTGCTGTTCTCAACATCATCATTGCCTGCTGATGTGCCACGATGCTGCTCCAATAGAC  
AGCAATAAGATTGTCTCTAATTTGAGCAGTAACATGATTGCAAGAGACCAAGTTTCACAGCTTGTA  
AAGTTCTGTATTTGGGATTCTTGCTTATTTTCCGCCTGTGTTTTTCTGAGAACTTATCCTGATGATC  
AATTGAATCCAGTAGTTTCTATGCTATTTGGTGGNGGATAAGCTACTGGAAGAACTATCATAA  
GGGAAAAATANAANGGAAACTGAATCATCTCTTGATTAAAANGGAATAAAGAAAGNAC

SEQ ID NO: 793 GTACAAATGTGCATTAACAATTCAGTGACGTANCTGTGGATCTCTGGATGGCT  
ATGTAAGCTGTGAGAAAGTCCCCCACTGGCTTTGCACTTGCTGCGCACCAGAGGTGACCATCCAG  
GCAGTATCAACCTGAAGAGGGGGTGAATCCCAGCAGCTGCTCGATGGGCTTAAACCGCCACTCGT  
CAGCCTCCAGCTCTTCTACCAAACAGCTAGTTTTTCCATCCGAGCAACTTGCTGATCATGTTTCAC  
CTGTTTAAGCGTGGATGCCACTTGATAGCTAAAAACAGATTTCGAGAGGAATCTCTCANAGCCAT  
CTTCTAACATCTCTCCAGCTCCTTTGGGGTAGGTGAGAAAGGGCTTTCCGTGGTGGGGCAAGGTCTG  
AGACACTGAAACCAGATCCAGTGACGAACCTTCCACTGACCCACAGCTGAGGAGGATGATGAT  
GAAGCAGCCATTTCTCTAGCAGAGTTCTTACTTCACTCAGCGGGCTTTAGCACTTTCCCTACA  
AGCCCCACCCCAAGGACCAGCCTGAGTGGCTGCGCTACACGACNAGGACCGCGGCGTGTTC  
CGCT

SEQ ID NO: 794 ACATCTGATTTACTGAATTTTAAAGTCTGGGATGTTAGTGGGGAAGAGGCGA  
GGTGAGCATTGCGTGACGCCGAGGACTAGGCGGGGGCGGGGACTGCACCTGGCTAGGCCACCCAC  
CCTGGGCAACTTGCCACGGACCCAGGGCAGTGAGTAGACAGGAGGTAGCCCGGGGTGAGA  
CCTCTACAGCAAGAAGATGGTGTGGTGTGGGCTCCCTGGAGAGTGTGCTCCCTGCGGCC  
TGGGAAGTGCTCCCTCACGACGGAAGGTTTCTGTGCTGCGGTCCCGGGGCTGATAGTGGCGG

TGGGCGGGTGGGGTCACGTGTCTCAAGGTCTGAATGCCCAGCTCTGCCCCATTCTCTGATTCC  
CAGTGGCTGCTAGCTGGACCCAGCTGGTGTCTGGGCATGANANGCAGGGCCACCCGTCC

SEQ ID NO: 795 ACCATAGTGATTGGGGAATTCCAGTTTAAAGCTCATAGGAATGTGCTGGCCT  
CCTTTNTTGAGTATTTTGGTGCGATCTACAGAAGCACTTCTGAGAACAAATGTCTTTCTTGATCAGA  
GTCAGGTGAAGGCTGATGGATTTCAGAACTGTTGGAGTTTATATACACAGGAACTTTAAATCTTG  
ACAGTTGGAATGTTAAAGAAATTCATCAGGCTGCTGACTATCTCAAAGTGGAAGAGGTGGTCACT  
AAATGCAAAATAAAGATGGAAGATTTTGTCTTTATTGCTAATCCTTCTTCTACAGAGATATCTAGT  
ATTACTGGAAACATTGAATTGAATCAACAGACTTGTCTTCTTACTCTGCGAGATTATAATAATCGA  
GAGAAATCAGAAGTATCTACAGATTTGATTGAGGCAAATCCTAAACAAGGCGCGTTAGCGAAAAA  
GTCATCTCAAATGA

SEQ ID NO: 796 ACGCGGGGAGAGTCCGTAAGGAGCAGCTTCCAGGATCCTGAGATCCGGAGC  
AGCNTNNGTCGGAGCGGCTCCTCAAGAGTTACTGATCTATGAAATGGCAGAGAATGGAAAAAATT  
GTGACCAGAGACGTGTAGCAATGAACAAGGAACATCATAATGGAAATTTACAGACCCCTCTTCA  
GTGAATGAAAAAGAGAGGAGGGAGCGGGAAGAAAGGCAGAATATTGTCTGTGGAGACAGCCGC  
TCATTACCTTGCAGTATTTTCTCTGGAATCCTTGAATCTTGAAGGAATGGACCTCAAAATTATG  
GCATCGTCAAAGCATTGTGGTGTCTTTTTTACTGCTGCTTGTCTGTGCTTATAGCTACGTATTATGTT  
GAAGGAGTGCATCAACAGTATGTGCAACGTATAGAGAAACAGTTTCTTTTGTATGCCTACTGGATA  
GGCTTAGGAATTTTGTCTTCTGTTGGGCTTGGAACAGGGCTGCACACCTTTCTGCTTATCTGGGTC  
CACATATAGCCTCAGTTACATTAGCTGCTTATGAATGCAATTCAGTTAATTTTCCGAACCCCT  
ATCCTGATCAGATTATTTGTCCAGATGAAGANGGCACTGAANGAACCATTTCTTTTGTGGAGTATC  
ATCTCAAAA

SEQ ID NO: 797 ACTTTTTTTTTTTTTTTTTTTTTTGGCAGGNGACTGAACTTTAATGCACCTAGG  
GATTGACCACAACATTTACAAATCAGTGTTACTGCATTGGTTTTTCAAAGATTAGCTACATTTATA  
AACTTTATTTTAAAAGTTTCACAAAGTTAAATTCCTTAACACTTCAACAGGAATCTAACTCTTTAC  
AGTCTATCAAAATGGTAATATTTACCTT

SEQ ID NO: 798 ACATCTATCATTATCACAACATGCTTATTTGATGAAGCTAAAGAAAAGCCAG  
AAGACTAATATGGCTGGATGAGAATAGCATTTTAAACATTTTTCAGCAAAATCTAGAAATTTGCT  
GTTTAGTTCCAAAACCTGAATTTTATACAAGTGCTATTATTCCAATAGTTTTTTTCAATTGTCCTAG  
TGATCTTGATCCATATACTGCATTCAATTATCCCAAAATAAAAGAGTTTTTNNATTAATNACCANG  
NGGTTTCAAAAGNNTAAATATATTTCAAAAACCAATGACNGTAAAAGNAGGGGTTCTAAATTTT  
TGCATGTCTTATTAAGCTTTGTTGNATAGACTAAAATTTTGGTTTCATAGACTCANAAATACTAAA  
TTGCCTATTGGGTTTTGGGGCTTCTTGCATCAGCTCCTGCAACTGTCCAATTTGTTTCATCACCGAAA  
GTCATTAAAGTTTTCTTCTGGGAGGAAGGTGCCAAAACATGCTNNTCCAGTTGAATTTTGTCTGTTTT  
CCGTTAGCCTGTTGCAATAACGCTGTGCATACACCTCCCATTCTTGCAAGAAACGCTGTGCCTCGT  
NAGAACCAACGGGNCTTATGTCTCCTAAATTTCTGCTTTTCAAGTC

SEQ ID NO: 799 ACAAACATATGTATCTGAAACACTTCTATTTGGCAATTTTATAACAAATCAAAT  
TTTAAAAAGAACAAAAGAGATTGCAGATTACTTCGCAGATACAGAATAAAGCAATTGATGAAGTG  
CTTAAGCAAAAAGAAAACAACAAAAAAGAAAACACACTGCTTTTCTTTTAAAAATAAAATCACA  
TTGCTATAGATCAAATGGATAATACCCTTATTAACAACCATTCAGATNGTTAATANAACAAGT  
GCTTTTATTTGCNCTTCACTTAATTTTATAAGACTCATTTTCATGTATATAGCTCTTTACCCCCATTG  
TTAACGAATAAAGTCCCTCATAATTTTACACTTTTAAATTTTTTAAAGCAAATGAGAAATGATTT  
ATGTATCGTGGAACCTTTCCCATTTTGGAAACCAAGGTTTTAATTCTATATTTTGTCTAATATTTT  
TTTAAAAAATTTAGTGTAATAAAGTCTGTTTTTATATCACTGTAGGTAAAGTGAAAACCTCCTCAA  
CCAGGAGTATTTTCTGCAGTTTCACTGCATATAACACATACTTTACAATGTCCTTCCAACAGTAA  
GGAAATAACTGATTTCTTGATCACTGTCAGAAATGAGTGCCATAATTCTATTATGGGTATAGGTCC

SEQ ID NO: 800 ACTTCAGCTGCTGATATGAAAATTAGATTATTTACTTCAGATCTTCAGGATAA  
AAATGAATATAAGGTTTTAGAGGGCCATACCGATTTCAATTAATGGTTTGGTGTGATCCCAAAGA  
AGGCCAAGAAATTGCAAGTGTGAGTGACGATCACACCTGCAGGATTTGGAACCTGGAAGGAGTGC  
AAACAGCTCATTTTGTCTTCTCCTGGCATGAGTGTGTGCTGCCTTCTTGAGGANCTTTTAAA  
CTTATGGGTGCAGAGAANAATGGAACAATGCCGTTTTATGATCTTTTGGCCCAACAGGCTATTTTA  
TCTCTTGAATCAGAACAAAGTGCCATTAATGTCAGCACACTGGTGCTTAAAAAACACCTTCAAAGTT  
GGAGCCGTTGCAGGAAATGATTGGTTAATTTGGGATATTACTCGGTCCAGTTATCCTCAAAATAAG  
AGACCTGTTACATGGATCGAGCCTGCTTATTCAGGTGGTCCACAATTAGTGAAAATCTGTTTGCA  
ACCACTGGTTATCCTGGCAAAATGGCAAGCCAGTTTCAAATTCATCATTTAGGACACCCCTCAGCCC  
ATCCTCATGGGTTCTGTAGCCGTTGGATCTGGACTGGCCTGGCATCGACTCTTCTTCTGTGTGT

SEQ ID NO: 801 ACTTTCTTTTTTTACAGTTTTTTTTTTTTTTTTACACACATATTATATACAACCTTTT  
 AAAATAACTGCACCTTCTCTAAAAGTGAGATATAAAATTGTGCAGCTATTTTAAAAGTTGTATATAA  
 TATGTGTGTAAAAAAAAGCTGTAAAAAAGAAAGGACAAAACAGGTTGTTTTGTTCTAGTTCTAAT  
 TTCTTAAAAACCACTACATGGTTCAAAATTGGAATAACATTTGGGGACACTNGGTTACCTCCAAAN  
 AAAAGGTTTTTAAAAAGAAAATGGGGTGGGATTGCCCATTTTGGATTAAATTTTGGCTTACAGTTCCC  
 ATAGCTGNTAGAGTCTGGNTTGGTTTTGNTTTTACTCTCAAAATCATAGTAAAGATCTCTCAGTCTC  
 CTGGCTAAAGATTGAAGGAAGGCAAAATCTATTTCTAATTATACATATATCAGTAAGGGATGATCTC  
 AACATAATAGTAATGTGTATCTTTTGGTATCCAGTTTATTTTGGCCTTCTAAAAAAGTGTCTCAT  
 AACACAGAACATTGCCATTTTGCTCTTGNAGGCCTCAAAATATGAAAAGCTATTAGTCCATAGAGCC  
 TAGGAAAAAGAATTGGATTAATGGGCCTTTATTTTGGNAACCCTTATAAATGCTGNANATATTA

SEQ ID NO: 802 ACTGGATTCTATAGGTCCACCATTAAAAAGCTGGTATGGGACACCAATTTATAC  
 ACATAAGGGTTAGATTAAAAATTTTAATTTTTTGGTTGATATTAACTGAAATTTATATAACTGAGG  
 TCTGAATCTTAAAAAAGTAAATGATAAAAAATTTAATATAGTTAACTGTTCACTGATATGTCTATT  
 CACTTCATCATAACCTATATATTTAATTA AAAATCAAATTATGAGTCTGCAAATCAGATGCTATCA  
 AGCAAATTGCCATCCAGGGTCCATAATTCTTTTTATATTTTTATCTCAGATGAATATATACGATTCA  
 GTAAATTTTAAATGTTCCAAATTGTTCTAAAAAATAAATTATCAAAAAGCTTCCAGTTAACAGTTG  
 GCTAATTCATTTGCCCCAACGAACCTACCTGTTTGTGTTGTGAGGTAGCATCAAAGACTATGATCT  
 TCTGTGACAGTAGTAGCCTTAATTCATACGCATTCCCTCTTCATAGGAAGAGTATGGCAACAAAA  
 GGGACAGATGAGTCCCTTTCATTAAATCATTGACTCCTGGGTTTTCATAGTATGTTAAATGCCGATT  
 TCAATTTTACAACAACAAAAAATCCAATATTTATCTGAAGGCATCGNTGTCCAAGGACANATTTA  
 ACACCTC

SEQ ID NO: 803 ACCTCACAACCAAAAAGCAGTTAACTATGCCTGGCATAACCACCCTGTCATGTG  
 GGCAGATCACTGTTCCCATGCTGGGTAAGAGTCTCCAAGAAGGGAAGCCCTTATGCAGTAATTCT  
 GTAACATATATAAATTTGGTAATCTAACATTAACTTTTTCGTGAGCAGTGAATATACACATGGTC  
 ATATGAAAACCTGCCCTGGAGACCGGGTTAATTATTAATAAACTAAAAGGGGAGAAATGCTGATA  
 GATAAAATTATGTCAATTCCAGGGTGTTCATGGAATAAAGAAACAGTAGCAGCTGCTTCAAAAG  
 TAGACTATGATCAGAAACCTCAGATGGTAACCTTTAAAAAATTGTGGAATCCAGAGTCTCAACCTA  
 AACCTACATAGGAGCTACAGCCAGGGAGTTCTTTCTAAGTTCTCAAGTGATTCTGATGATTGGCC  
 AAGCTGGAACTTCTGAACTGGGGGAGAAAGCCAACCTANGTAAATGTAGAGACTTTTAGAATTC  
 AACATTTGATTATCATTTTACTCCACATTTTCCCCACTGTCTCACCTTCTCTATTTTCTAACAGTGGA  
 CAATAAAAGGGCTTGGCAAGTGAANGGAAAAAGTCANTAGAGTCCT

SEQ ID NO: 804 ACTGGGAACAGGTGCTTGCCTTGCTATGGCCACGGTTTGAAGTATCCTGGA  
 GATGAATGTTTCAAGCGTCCGAAGCACTGACCCCGAGCGCCTAGGGGGGTTGGATACTCGGCCCC  
 ACTATATCACACGCGCTATGCAGAGTTCTCCTCCGCTCTTGTGAGTATCAACCAGACAATTCCTA  
 ATGAACGGACCATTGCAATTGCTGGGACAGCTGCAGGTTGGAGGTGGAGAATTTTGTCTCCGAGTG  
 GCAGCTGAGTTCTCCTCAAGGAAGGAGCAGCTTGTGTTTCTGATCAACAACATATGACATGATGCTG  
 GGTGTGCTGATGGAGCGGGCTGCATATGACAGCAAAGAGGTTGAGAGCTTCCAGCAGCTGCTCAA  
 TGCTCGGACACAGGAATTCATTGAAGAGTTGCTGTCTCCCCCTTTTGGGGGTTTAGTGGCATTTGT  
 NAAGGAGGCTGAGGCTTTGATTGACGTGGACAGGCTTGAAGGCTTCCAGGGGGAAGAAGCCCGG  
 GTAACCTCANCTGATCCGTGGCTTTGGTAGTTTCTGNAAATCATCANTNGAATCTCTGATCAGGATG  
 TAATNGCGAGTTTTACCAAATCCAGAAATGGCACCCANTNTNTTTCAGGGANCGCTAACCCANN  
 TGAACCANGCTNTATAAT

SEQ ID NO: 805 ACTGGATTCTATAGGTCCACCATTAAAAAGCTGGTATGGGACACCAATTTATAC  
 ACATAAGGGTTAGATTAAAAATTTTAATTTTTTGGTTGATATTAACTGAAATTTATATAACTGAGG  
 TCTGAATCTTAAAAAAGTAAATGATAAAAAATTTAATATAGTTAACTGTTCACTGATATGTCTATT  
 CACTTCATCATAACCTATATATTTAATTA AAAATCAAATTATGAGTCTGCAAATCAGATGCTATCA  
 AGCAAATTGCCATCCAGGGTCCATAATTCTTTTTATATTTTTATCTCAGATGAATATATACGATTCA  
 GTAAATTTTAAATGTTCCAAATTGTTCTAAAAAATAAATTATCAAAAAGCTTCCAGTTAACAGT  
 TGGCTAATTCATTTGCCCCAACGAACCTACCTGTTTGTGTTGTGAGGTAGCATCAAAGACTATGAT  
 CTTCTGTGACAGTAGTAGCCTTAATTCATACCCATTCCCTCTTCATAGGAAGAGTATGGGACAACA  
 AAAAGGGACAGATGAAGTCCCTTTCATTAAATCATTGACTCCCTGGNGTTTCATAGTATGGTAAATG  
 CCTGATTTCAATTTTACACCACCAAAAAATACCATTTTATCTGAAGGCANTNGNTGTTCCANNA  
 NCANATTTA

SEQ ID NO: 806 ACTATTGCTATTAGGGGGTCTGTTTTATAAAATATTTTCTTATCATACTTTTATT  
 ATAACTTTTTTAGTATGAAATTTGCTTCAACTGTTACAAACAGAATCATTTTCTATGGGGTCCCT  
 CCACATAAGGAAGTTATTCCTGTAATTACTATTTTAAATAGTCTTCTTAACTGTGGGAAAACCTTN  
 CCCTNCCNCAGCACGCACTCTTACTCTCTGNTATGANGCTNAATGCTNTCCAANGCACTGGTT



CAGTCANCATTCTGCCCAT

SEQ ID NO: 807 ACCATATGGTAATGCTGCCTGTCTTTCTGAGGTTGACTTTTATGCCATGTCTTT  
CCTAAGTGTGTAAGAATTTTTCTGTTTGCTTCACATTTGACTGAGAATCATTCTAGGGTTTGATTGA  
GCCCCTGTCCTGTGCCACTAAAGGAACTCGAACTTTTCATCACTTAGAGATTTTACAGAGGGGAATGG  
AAAAACAGTTCTAATCAATAAGCAAGCAATTCAAGAAAAATAGAATTAATCAGGCAATGGCTGCA  
ACATGTCCTATCTTTAATCTATTTTCTTATTAAGCTTGGACATTGACAATAGAACCAGAAGCTTGTA  
GCTGGATCAAAACATTCTCCATAGGCCTGGAGTTTCATGAGGGTCTATTCTTTTGTGTGTTGTTT  
TGGTTTTTTGTTTTTTGTGGGTTTTTTTTTTTTTTTTTTGAGACGGAGTCTTGTCTGTTGCCAGGCTGG  
AGTGCAATGGTGCAGTCTTGGTTCAGTCAACCTCTGCCTCCCAGGTTCAAACAATTCTCCTGCCT  
CANCCGCCAAGTACTGGGATTACAGGTGCATGCCACGATGCCTGGCTATTTTTTGNATTTTAAGTA  
NAGGNGGGGTTTCACCATGTNGGCCAG

SEQ ID NO: 808 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGNTTTTTTTTTTTTTTTTTTTTTTNA  
ANACNCAAAANCAAAACAAGTTTATNTTAAAAATACCAAGTATCACATCCATANCCTATTCTTAA  
TTNTAATTACTTCTTAAAGCAACNTTTGNCTTTGTTTACAAANNCAGCNCAGGTGATAAAACCTA  
NTNTCTCAAACCTGATTGCTTACANITTTATAATACAAACNTTTNTATATGTCTCAAATCACATTTTAG  
GCAGTTTAACTTAAATGCATANACAATTACATACATTAATTTGNCATTACTGGCAGTGAAAAAT  
CACAATTTCAAGAAGAAAATTACAACAACTCATAGAGCACACATAANAAGTCAAGTTCAGTTAA  
TGGCAACAAGTNANTATCTGCTGATCANTGACAGGCNNTGAAAGTGCCTTTATCCAGGGGTAAAG  
TGAAAAGGCAAACTNTCCTGTTCTTTGCAACTTGAATTGTGCNCTNATCAACACTACAATNTTTGA  
AAAACNTNAAATTCTTAAAAAANGCTTGC

SEQ ID NO: 809 ACGCGGGGTGGGGGGGGTCTGGTCTTTGGCTTCTCGACTCGGTCCTGTTTCG  
ACAGCGAACATGTCGCGCCTGTGAGAAATAGGAAGGTTGTTGATTACTCACAGTTTCAGGAATC  
TGATGATGTANATGAAGATTATGGAAGAGATTGGGGCCCTCCCACTAAGAAAATTCAATCATCTC  
CCCGAGAAGCTAAAAATAAGAGGCGATCTGGAAAGAATTCACAGGAAGATAGTGAGGACTCAGA  
AGACAAAGATGTGAANACCAAGAAGGATGATTCTCACTCAGCAGAGGATAGTGAAAGATGAAAAA  
GAAGATCATAAAAATGTGCGCCAACAACGGCAGGCNGCATCTAAAGCAGCTTCTAAACAGAGAG  
AGATGCTCATGGAAGATGTGGGCAGTGAGGAAGAACAAGAAGAGGAGGATGAGGCACCATTTCCA  
GGAGAATTCCCGGCATGCNATGAANATTTCTTAATTGGAAGATGATGACGATAGTNACTATGCA  
GTTCCGAAAAA

SEQ ID NO: 810 ACCGCGGGATTAAAGCATTTGTTCCAATAAAATAAATAGAGGGGAAACTTG  
GATGCTAAAATTACATGAATAGGGAATCTTCCTGGCACTTAAGTGGTTCTATGGTATTGAAAAAT  
GGATGTTCCCAGAAAGAATTACTTTTTCTCTTATTTTTACTGCCATTGNCGACCTATTGNGGGAC  
ATTTTTATATATTGAATCTGAGTTCTTTTTGACTTTTTTTTTTCCCAATCCAACAGCATNCTTTTTT  
TAAAGAGAGAATTAGAAAATATTAATCCTGCATGNAATATATCTGNTGCATCTTAATTGGACC  
AACNTCCCATTATTTATCTTAAACTATACCGTTACCTCTTAATTCCATCCAAAGANATACAGTT  
TGAAACAGAAGTGACCTTGCCGCGACCCGCTAAGGGGAAATTTCCCACTGGCGGGGCGNTA  
CTANTG

SEQ ID NO: 811 ACAGAAATGGTATTTGTGTATGTGTGTGGGCTTANAGATTACAAAGTAAATATT  
CCTTTGGTGAAGGAATTTCAATAAAAACATCTATCAAGTGTGAGCGGTGAGTGTGTTACACCACA  
GAAATTGGCAAATTGACAAATCAGAGTTTGATTTTGTNTNNGTTTTACTTTCCATAAAGTTTCGT  
TTACCAGCATCCACTAAAGATTNCGGTTTACAAATAAAANCCNTCTCGNTTTGAGCANTACTATN  
CTACTTTTAAANATTNTCNT

SEQ ID NO: 812 ACTGAAAAGTTGCCACTTTTTATTTAGTAAGAAAAACAAACATTCTGGCTCACT  
AGAGTTCAGAAAAGTAATAATTTGAGCCAAAGGAATTTGAATTAAGAAAATAGAACTAGGTTTC  
ATGTATTTAAAAAATAGGAAATAAAATAGAACTCAAATGCCATGAAGTTATCTTCCTCTTCCTTA  
TATCCCCTAAGTTTGGGTTGCAAATAACTTTCCAATTCCTAATAACCTAAATTATTTTGAATAGTG  
TTTTCAAGTGAAATGATGAATGTTTGAATGTTTGTGTTGGTAATCAACATAACANTGCTNACAAAATT  
CCCCTTAGATTTTTAACTTTTAAAGTCAACGTTGTTGATAATTTGATAATTTAAATGCCAC  
ACATACACACACACACCCCCACACACACAGACATGTTAATAAACCTGATAGGATGGAGTGAGCA  
AAATTGTTTTCANGGAAGATGGCCCATTTTANAAAAAGACTTGCCAGATGATCTTCCATATTTGC  
TCCCCCAATGGNATTTTTCNCTCATGCCAATCCTCCTCTTGTACCTGTGGCTATGGTTAAATNC  
CTTTAACAAGCNTTCATTGGGTCCNAGTATTTCANANACTTTTTAAAAAGCTTGGGCATGACTTAN  
CGTTTGTGCTGT

SEQ ID NO: 813 ACAAATATGTATCTGAAACACTTCTATTTGGCAATTTTATAACAAATCAAAT  
TTTAAAAAGAACAAAAGAGATTGCAGATTACTTCGCAGATACAGAATAAAGCAATTGATGAAGTG



CTTAAGCAAAAGAAAACACATTNAANNNGAAAACACACTGCTTTTCTTTTAAAAATAAAATCACA  
TTGCTATAGATCAAATGGATAATACCCTTATTAAACAACCATTCAGAATGTCTTATAANTAGCAT  
GTGCTTTTATTTGCACTTCACCTAANTTATAAGACTCATTTTTCATGTATATAGCTCTTACCCCCA  
ATTGGTTAACGAATAAAGNCCCTCATAATTTTACACTTTNAAATTTTAAAGCAAATGAGANATG  
ATTTANTGTTCGTGGAACCTTTTCCATTTTGGAACCAAGGTGNTAATTCNATATGNTTGGCTTAA  
NATNCCTTAAAAAATGTAGTGT

SEQ ID NO: 814 ACTATATAGTTTTTAAAGAATGTTGTCCCACCAACTATTCATCCAAGCAAAGA  
ATTGTAACATATNAATAAAGTCTCAGTTACACTTTTGCTTTATCACAATAATTCATTGTAGAGCAT  
TGCGCAGGTCCAAGAATANAGCTGCTCAAAATCTTTGTGGNAGTTTCTTTAGTTTTTGTAAACCTT  
GAGGCATATGTTCCAGAGAACAGGGATATTTGACTGGTCCAGNGACCTNNGGTGATCATAGTCATA  
ATTGAAGAGATGCCTATGGGCATGCTTAAATCNCAATTGCCAANTGATATNGNTGNTGNATTATT  
TCACNTTCTTGGATCTATGNATGTAGTTGGNATAACAAATATTTAAATAGNTNTTATTTTGTATGCC  
ATNAAAAAATCATTACTCTGGGCNTTTTNTCCCNNTACTGTNGTTTCCCAGAATCTTTTAAAN  
AATTCAT

SEQ ID NO: 815 ACATTTTTGAATAGACCTCAAAAATACTTCATTCTGCTGCTGTTTCAGTTGGCT  
TTTTAAACCTGTCTGCAGTAGGACACTGAAAACAGCAAGAACTTCGGGGTGAACACCCGCTGATC  
CTTTAAACAAGGATTTCTGGCAGGAAACTCACAAAAGGAGAACTGAAAATTTAGACATACAGTTG  
GCCATTGTAAAAAACATCAGTTCCCTCTCATACATTCCAAGTAAACCAAGTAAATAAGTGTGGG  
AAGTACACTTGCATAAAAGAATTTAAGGGAGGTAGCTCTTTCTGTTCTGCCATTTCCCAACATTC  
CTGGGGGGAAGGAGACTCAATGAGTAATACTATTTCACTGAGCCCAAGATGGAAACTTGGTTTG  
ACCTAAACATCTGATTAATATAGCTAGCTGATTTCTTAAAAATTCGTTGCATTGAANGATATTTTG  
CATGTCTGTAAACACCTGGCAATACTTGNTTGNATTGGATTCTGGATATTTCTTGCAGCTTGACTACG  
TGTAATTTGGGCCAGATCAGCTTTGCAGTAAATTATGCTGCATCCTCGTGGCAAAATCTTNATT  
CTTAGNGAATGGTACCAACCCCTTTATTGCTGGCTTAAGAAAGTGAAAGATTGGTGTATTTCTTT  
TAAACATTTTCAATCA

SEQ ID NO: 816 ACTGTTGGTTAAATGACAATTTATGTGGATTTTGCATGTAATACACAGTGAGA  
CACAGTAATTTTATCTAAATTACAGTGCAGTTTGTAAATCTATTAATACTGACTCAGTGTCTGCCT  
TTAAATATAAATGATATGTTGAAAACCTTAAGGAAGCAAATGCTACATATATGCAATATAAAATAG  
TAATGTGATGCTGATGCTGTTAACCAAGGGCAGAATAAATAAGCAAAATGCCAAAAGGGGTCTT  
AATTGAAATGAAAATTTAATTTTGTTTTTTAAATATTGNTTATCTTTATTTATTTTGNNGTAATATA  
GTAAGTTTTTTTANAAGACAATTTTCATAAATTGATAAATTATAGTTTGTGTTGTAAGANAGGTTGC  
TCTTAAAAGATGTAAATAGATGACAAACGATGTAAATAATTTTGGNACAGCTTAAATGTTATAC  
NTGAAACCACNTCATGAAAGCNGAATTTGGTTGNGTTTGTNTTTNCTCTATTTTGATNNGGGCA  
TACT

SEQ ID NO: 817 ACCCCAACCTTTGCTGGACCTCATGCAGCTTTAGCTAATAAAAAGTTTCTTTAAG  
GCAGATAAAGTTACAATGCTGTGGAATAAAAAAGCTACTGCTGTGTTGGTAATAGCTAGCACAGA  
TGTTGACAAGACAGGAGCTTCCTACTATGGAGAACAACTCTACACTACATTGCAACAAATGGAG  
AAAGTGCTGTAGTGCAATTACCAAAAAATGGCCCCATTTATGATGTAGTTTGGAAATCTAGTTCTA  
CTGAGTTTTGTGCTGTATATGGTTTTATGCCTGCCAAAGCGACAATTTCAACTTGAAATGTGATCC  
TGATTTTGACTTTGGAACCTGGTCCTCGTAATGCAGCCTACTATAGCCCTCATGGACATATATTAGT  
ATTAGCTGGATTTGGAATCTGANGGGCAAATGGAAGTGTGGATGTGAAAACATACAACTTATTTTC  
TAAACCGGGGCTNTGATCTAATATTTGTTGNGCCCGATGGNGACANTTTTACAGCTCTGGCTCCA  
GTTACGGTATATGATACAAATNG

SEQ ID NO: 818 ACAGAATGGTATTTGTGTATGTGTGTGGGCTTAGAGATTCACAAGTAAATATT  
CCTTTGGTGAAGGAATTTCAATAAAAACATCTATCAAGTGTGAGCGGTGAGTGTGTTTACACCACA  
GAAATTGGCAAATTGACAAATCAGAGTTTGTGTTTTGTTTTTGTGTTTTTACTTTCCATAAAGTTG  
TTTACCAGCATACCACTAGAGATTTCCGTTTACAAATAAAAAGCCATCTTGGTTTGGCAAGACTAT  
GCAACTATGAAAATGTTTCGTTTAAAAAATCTTCATGATCCTTTTGTAAATACAAGGTGGTTGCCA  
AGCTTGTTAGTTTGTGTTATTTTATTGATAGATGTAATAATTATTGTAACCTTATTGGATAAAGTN  
TTCAAAAGAACANNAGCTTACAATGAGGGAGGNTTNGATTNTTGCTANGTGAAAATTGCAATTCA  
AANNTCTGCTNCTCTCTGGNATGCANAAGAGAANNAATCCTTTAAATTTGACTGATTTGAAATCA  
GACTGNTTCNATTCTCGANCGNNACTTACTCCTTTTTTTTTGGANTAAAC

SEQ ID NO: 819 ACATTGGAGAAGCTGTGCAGCAGCATCCTTTTCTGTGGTGGGCAGGGCAGGA  
GATGAACCATAGGAGCCAAAAGTCAGACAAACAGAAGAAGGCACACCAAGCCTGAACCCCTCCGG  
ACAACAGCAGAGTTACCAGCTGAGGGATGTCCCTGGAGGTTTCTGACCCATGAGAGGCCCCCTCA  
CCCTCCTTACCCTCCTCCTACCACCAAGCTCTCCGGCAGTCATGGACTTATTCCTCCCATTTTAC

TGGACACCTGNTTTCAGGTCANTCTGTACCTGTATCTTGGNGGATCTCTACCCCTACTGCCAGAA  
GGTGGGATGTGGNTTGGATGGACAGCCANATTCTACTACATGGCANTTTGACAAACACCTCCTTG  
AAACCATNTNTNTTTGTNGGTANNTCTTTAANAANCTTTNTTTTCGNTGTCNGCTANCTNTGGNTT  
GATNATTTTCCACTG

SEQ ID NO: 820 ACTCAGGGGCATCATGTTGCTGCAGAGGCCACACTTTCCAGAAGTTTCTCCT  
CGCTGTGATCCTCGCACACCGGGGGCACTCGGAGGACTGGAAGCACTGTTTGTGAAAGCAAGCCC  
TGCACGCTGAACATCTTCTACATGTTGCTGTCTGAAATGGGAAGATGACAGTCGTATTCTGGCAAA  
ATTACAAATAAAGCCCTTTCCTTGACACAGCTCACAGCCAGCCACATGTGCAAGGGAAGCTTTC  
AGAATGTCTTGAGTAAGGGTGCCAGCAGCCCTTTCTTGATCCTGACCAGGTCCTCAAGGGAGAA  
CAGGTGGAGCTCATCAGCAAGTGCCCGGCACCTGCTCGAACTCCTTTAATGCCTGTAGCAAACCT  
ACAGGTCTTCAACAACTTCTTGNTTGGAAAAGTGTCTGAATTCCTCACTTTGCCACTCTTGNTT  
CCATACAGGCTTGGCGATGCTNACAAATTAATAATGNTTGTCCATTTCTGCAACACTTTTGGAAAA  
TTGTNACGAAAA

SEQ ID NO: 821 ACAAATCTTTGGCCTTTCTCTTGACATTTTCGTATGTCAAAAAGCAAAAACC  
TTCATGTATTTCAATCTAGTGATTACTTTTGCACCATAATTTGTTTTTACACCACAAAAGGAGGC  
ACTTTCAGTATCTGTAAAAGGTATTTAATCCTAAAACATACTTACCTAGAGAATAATTAACACAGA  
ATTCAATACAATCTAGTATCTATTAGGAAATTAAGAGTTATCACTTCTAAAAGTCATTTGAAAGTC  
AATGATGTTATCTGGTCAATGGCAGGAAATGGGAAGTGGAAACCAATATAANACTTATGGGGATT  
CCTCACGGAGACAAAAAAGATATTCCTTTATGTTGTTTAAAAAGTGGCAGCTGCTCTTTCTTTAT  
TCCATTTTAATCAATGAGTATTGATTCAGGTTTCCTTCTATTTTTCCTTATGATAAGGTTCTTACA  
GGAGCTTATTCAACAACCAATAGCATAGAAAACTACTGGATTCAATTGATCATCAGGGAATAAG  
TTCTCAAAAAAACACAGGCNGGAAAAATAAGCAAGAATCCCAATTCAGAACTTTACAAGCTGNGG  
AACTTGGTCTCTTGCATCATGGTACTGGTCAAATTAANANACAATCTTATTGNTGNCTATTGGAGC  
NCCATCGNGGCCCATCAGCAG

SEQ ID NO: 822 ACTTTTTTTTTTTTTTTTTTTTTTTTCAAAATGAGACTTGGAGTTTAATTAAAAA  
CAGAACAGGGATNCNTTAAACAAACAAACAAAAATTACTTTCTGATTATCAATTTTTTTGANAC  
TCAAAGCATCCCCAAAACATTGGANATCCAGCTTATTCTGANACATCAACCATCACAAAAGGTTT  
TCACTCTGAACTATTACATTTTNGTGGCANAAAACANAACAAAGTTCTGCANACATCCTTCCTCT  
CTTTCTAAAATATATTACAAACAGGGTNTTTTCATAGTCAAAAAGAAAAACAAACCAGGTTTCTTT  
TTTGGCCAAATGGGCCTGTACTCTCCCTGGGATCTGATTTCTTAATAAAAAAGTTCAGGGCACC  
AAATCCAACCAGAAATTCCCAGNACCCAGNGGCTACTTAACTATGAGGGGATGGATGCTTTTGT  
CTTTTCTATTGAGGGGAATCATTTCTCCCGGATTTNTGCTGCTCAACAGCCCCAGGACAGGTANG  
TGGGAANGNNGGTGAAATGCAAAACCAAAAGGGTCCAAAAAANAATGAGGGTTTTTGNACA  
AACCATANCAAGNAAAATNGGCCANTTTNCAAAACNCCCCTTNAAACTCAAACTNCCCNCC  
CAAACTTNGNGGGGAAANGG

SEQ ID NO: 823 ACTATCTCCTGCAGCTCAATATAAACAACGATGTATTTCTTTGTAGCAGTATG  
TGGAGTCAGAATGGGTTTTCCAATCTGATTTAATCCCTGAGCAGCTATGTTAAAAGTTAACTTTTT  
AAAAAAGCCCAATAAAAACTAAATACTGTATTAATACTCTACCATAATGTTACAGGGATAAACT  
TATTTCTGTACCCGATTTTATTTCCAGTTTTTCATCCGAATCTACTGTGGAATGGGATAATTTTGCTTT  
TGTTTCATTGCCAGGAATTAATCCTCAGCCTCCGAGTAGCTGGGGACTATNAGCGCTTACTAGCC  
AGCTTCAGGGCTAGAATAGCCTTTCTGTGCAAGTGGAAAAAACCTCANGNATTTCCCAAATCGGT  
AGCTGCAGAGGGGCTCTATCCTACCAGAGATCCCGCGTACGACGTCNCATCTGGGAAAGAATCTA  
AAGTTGGTGANCTTCACCAACCTGGAGCATTCAATGGGGAATGCACTCATACNNACTGGGCATGA  
ACCNCTGGGANACATGACCAGTGAANAAAGTGGNTGNCTTTGATGANTTNCCTGGAGAGTTCCC  
ACCCCGGGGCAGAGAAAATATCACATTTTAGGCCAACCCCTAATCGGATNTTGGCCTGAATCTGN  
GGGACTGGGANANAANAANGG

SEQ ID NO: 824 ACTTGGGGATATAAA  
CTATTTATTAACANACAAGGNCTACANACTTATTTNTTTTTGGACACACCCACGGNGCGGCCACGG  
GGGCCAGNGGNTTGGGGNGCTGGCCTCGGACNCAAAGGCCCAAAAGNGACNCANCCCTTTAT  
GGGCCCGAATCTTNTNAGNCGCTCCAGGNNTTACGGAGCTTGNTGNCCAAACCATTGGCTANG  
ACCTGGNTGNATTTTCCATNCTTTANANCCTTNTGGCTGGGCAAAAANCAAATCTGGGNNTTTTTG  
TCCCGGGTTATNTTACCTTANTTGGGNAANCCAACAAAGGGAATGNTNATAAGGGAAAGGNTAA  
AAAAAGNTAAAGGGAAC

SEQ ID NO: 825 ACGGACTTGTGCAAAATGAAGAGAATTATTAGCCTGTTCTCTAAAAGCGATA  
AACTCTGGGATCTTCTCAAATGCGCCATATTTGTAAGCTTGAATAATATATCTGAGGTATCTTTCT  
GGTTGGAGTGAAAAACCTGAGTGCGAAGTTACAGGATTGGGACGCAGCAGCATACTGACCTAGA

GATTCAGCATATCGGGTCAAAAAGATAACCAATGGTATCATGCTGGATATGCTTAGCATCGAGGCT  
GGAATAAAGATCCACCACTGGTTCAAATGCCCCAGCATACAGTAGATTGGAACAAGCAGCAATTT  
GAACTGAGCATTGGAAGGGCTTTTGGGTAAATCCCTCTTCCAACAAAGTCAGGGGCTGCCACAC  
AAGTGGTCTCATCACCTGTTTCCCTCCATACATTCAATAAGCCCTTGGACAGCAAGGAGACAGTAA  
TANTCANAAAATTGCAATTCCTTGTCTTCAAACAAGGTTTCCCAAATTCCAATCCATGCTGGT

SEQ ID NO: 826      ACGCGGGGCCCTTACGCGCCGGAGAGATGGCGGAGTTGGACATCGGGCAG  
CACTGCCAGGTGGAGCATTGCCGGCAGCGAGATTTTCTTCCATTTGTGTGTGATGATTGTTTCAGGA  
ATATTTTGCCTTGAACACAGAAGCAGGGAGTCTCATGGTTGTCCTGAGGTGACTGTAATCAATGAG  
AGACTGAAGACAGATCAACATACATCTTACCCATGCTCTTCAAAGACTGTGCTGAGAGAGAACT  
TGTGGCAGTTATATGTCCTTATTGTGAGAAAGAATTTTTTGCCTGAAACACCCGNCATTNANTCAN  
ATCNTGGANGTGGAAAAATTGNAAATTCCCAAAGCCTTGAATGGGNTGCCACTNNGAAACTTGTT  
AAAANACAATTATTTGATTTCCAAGNCAGGGAGAAANACANCAANTTAACCATGGGAAANGTGCC  
NAAAAANTNGTGAAAAACAANTGNNNAGGGTTGCATTGANTNAAATTAAGATGCATGCNTGGTG  
GGNGTTAATCTTTTACCCAGACAGAAAGAATTTCTTTNNNGGTTTTTTTAACTTAAAGGGAGC  
AAAGAAAAAANCAAACCTATGTTCTTTTGGCCCNNTTNGNGCNTTTGGNAAGGCCNTTAACTTTNCC  
GGTTTTTTTGCAAGGCTTTAAAT

SEQ ID NO: 827      ACTCCAGGAAGATGCCATCTTGCACTCAGAAGATAGTTTAAAGGAAGATGGCA  
ATAATAACAACACATCTTCAATACCAGCAAGAAGCTATTGAGAAGAATGTTGAACAGTCATCGGA  
TCTACAGGACCAGTTGAATCATCTGTTGAAATAGAATGACATTAAGTCAAGGAGATACGTGTTTT  
ATTTGTGATAGCAAATTCCTAAATGAACATTAGGCAAGTGGTATCATTATCAGGCCAGCTGCAGCC  
TCTTGCCTTGACCTGCATTCCTAGAATTTCTTTGTTGCTGNAATTTCTTGGATTAAGTGACCTTGACT  
TTCATTTTGNAAATTTTGCTAATCATCAACAAATTCATCTGCATGACGTTACTGCCAAATATGAAGG  
GCAGTTGAATTATATGAGTGATTGTGGCAGANGTTTGTGCCATGGNGAAAACCTTIGATGGTTGNC  
TGGGGTCATTGGATCCATCTTTTTAAATGACCTTACCATGAGTCTGGTTGNCAAACCTAAATATCTT  
TGGTTGAATTTAAATGGGACTCTTATTGGTGNAGTTCANGNCTTCATTGCTTAAAAAATTGNAG  
AAATCTGCCATAAGAAAATTTTGTTCCTGCNGGAATAAAGAGGAAGTAACAGGGAATCCCATAT  
TGGTCATATTGGGTNTTGG

SEQ ID NO: 828      ACTTTTTTTTTTTTTTTTTTTTTTTTTTCTNAGTATTCAACACTTTAATATTTATG  
GNGTATCACATAAAAAACAAAGTCATATACTTTTGCAATTAATCAAAAAATAGCAAATCCATATAA  
TGGCAAAATCAGGAAAAAAATNTAGTATTTCCACAAAATACATAATGTCTTACAGATGATTATGT  
GAACTTTAAATGTCTGCAGCCCTACANAGCTTTTGTGCCAATTGAAAAACAAAAAATCCCAAC  
ACAGGATGTTCAAAAAGCCTAATTCATAAAAAGACAATTTATCCATGTTTAAATATAGNGTTTTTT  
AGGATGGTAACATAAGTCATGCAACAGCTCTGTAAAAACAAAACAAAACAAAGAACTACGATGTC  
GGCTGCGGGTTAAATAAAAAGGAAAACCNENCATACAAAAAATGTAAGGAATGGTTAGTGGT  
GCTGCCAATTAAAAAATACTGGAAATNATTTTACCCCCAAAAGTGATTNTGGAAAAACTGTT  
TGGAATCTNAACATNGGACTTGGGTTGNAGNCATCTTTTGGGAAANTATAAGTGAAAGNGGTTGG  
GGACCTCTGNGGTTCCATTTTAAAAAATAATTGGT

SEQ ID NO: 829      GGTACTGCTAGCTGGAAGACACATAGTGGATCTGTATGGCGTGTGACATGGG  
CCCATCCTGAATTTGGGCAGGTTTTGGCTTCTCTTTTGACCGAACAGCTGCTGTATGGGAAG  
AAATAGTAGGAGAATCAAATGATAAACTGCGAGGACAGAGCCACTGGGTTAAAGGACAACCTCT  
GGTGGATAGCAGAACATCTGTTACTGATGTGAAGTTTGCTCCCAAGCACATGGGTCTTATGTTAGC  
AACCTGTTCCGCAGATGGTATAGTAAGAATCTATGAGGCACCAGATGTTATGAATCTCAGCCAGT  
GGTCTTTGCAGCATGAGATCTCATGTAAGCTAAGCTGTAGTTGTATTTCTTGGAACCCCTCAAGCT  
CTCGTGCTCATTCCCCCATGATCGCCCGTAGGAAGTGATGACAGTAGCCCCAACGCAATGGCCAA  
GGTTCAGATTTTTGAATATAATGAAAACACCAGGAAATATGCCAAAAGCTGAAACTCTTATGACA  
AGTCACTGATCCTGTTTCAATGATATTTGCATTGCTCCCAAATTTGGGGAAGAACTTTNNNTNTNT  
ACCATTACCACCAAAGATNTGANAATTTTCNTTTAAGCCG

SEQ ID NO: 830      GGTNCAAGTGATTGTGACAAATGACGTAAAAATGGCATTTCATGATGTCTGAA  
ACAAGCCTAAATAGAATTCAAGATTAGACTAAATGATTTTCACAAAGCACATTCAAGGTTTTACAT  
TCTATGATTGAAAAAATTTTTGAAAACCTTTTATTTTCAATCTTTCTGTAGGATTTTGCTACAAA  
TAACTTTGGGAATGAATAAAGTGGAAATGGTAACTTTCCAGTGGTTCAGAATTGAATTAGACTTCTT  
GTGACTGTGATGTTTGGTTTCCATTGAAATATATGAAGTGAGATGTCATATCCTGAATATAGTTTG  
TCTTCCCCAATTACTTGATAGCATGCTGTGCAGCCAGTAAAGATTAAGAACAGAGTTTCTCTAAAT  
TCCTCCGATTATTTCCACTAAGGCACATTAATACTTAATTTTGGGAAACCCAGACATCACAGATTT  
CTCCATGAAGTCCTAAATCTTCTTAAAGTCAGAAATANGNATCTTAGTACTGACAGTATTCAGGN  
TTTTTCTNCCTTGGGGGATATGTCATTNCATCAGNGAAAAAATTTNTTCCNCGGATAANAAA

AGGGTTCTNGNGATACATATTNNCAATCCT

SEQ ID NO: 831 GGTACCTTCTTTAGTAGAGACTGGGTTTCACCATGTTGGCCAGGATGGTCTCT  
ATCTCCTGACCTTGTGATCTGCCTGCCTCAGCTTCCCAAAGTGCTGAGATGACAGGTGTGAGCCAT  
CAGACCCAGCATTTTTTTTTTAATTTAAATTTAAATTTTTTTCATTTTTTTTGAGAGGTTTTTTTGT  
TGTTTTGTGTT  
AGTGCACTGGCATGATCTCTGCAACCTCTACCTCCCAGGTTCAAGCAATTCTTGTGCCTCAGCCTC  
CCAAGTAACTGGGACTACAGGTGCACGCTACCACACCTGGCTGATTTTTTTTATGTTTTAGTAGAG  
ACAGGGTTTCAACCATGTTGCCCAGGTTGGTCTCAAACCTCCTGAGCTCAGGCAATCCACCCGCTT  
GGCCTCCAAAGTGCTAGGATTACAGGTATGAGCCCCCAGGCTATTTTTCTTTCTGTTTTAATTT  
AAAGTGGGGGGGGCTAATTGGTATCCTGNTGACTCGAACTCCGGACTAACGAATCTTGGTTCAN  
CCT

SEQ ID NO: 832 GGTACCGGAAGAAGCAGCTGGCAAAGCAGCTCCCTGCACATGACCAGGACC  
CTTCAAAGTGCCATGAGTTGTCTCCAGAGAGGTGAAGGAGATGGAGCAGTTTGTGAAGAAATAT  
AAGAGCGAAGCTCTGGGAGTAGGAGATGTCAAACCTCCCTGTGAGATGGATGCCCAAGGCCCCAA  
ACAAATGAACATTCTGGAGGGGATAGAAGCACCCAGCAGCAGTGGGGGCCATGGAGGACAAA  
TCTGCTGAGCACAAAAGAACTCAATATTCTGCTATTGCTGCAAACCTGAGTATGAAAGAAGGTGA  
CCCAGCCATCTATGCCGAAAGGGCTGGCTATGATAAACTGTGGCACCCACTTGTTTTGTCTGCAGC  
ACCTGCCATGAACCTCCTGGTTGACATGATTTATTTTGGGAAGAATGAGAAGCTATACTGTGGCAGA  
CATTACTGTGACAGCGAGAAACCCGATGTGCTGGCTGTGACGAGCTGATATTCAGCAATGAGTA  
TACCCCGCCGAAANCAAAATTGGCCCTGNAACACTTCTGTGNTTGAAGTNGATAACANTCTTA  
CTGGGGAGAAATCCTGATGGCAATGACAAC

SEQ ID NO: 833 GGTACGCGGGGGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCGCGGTGGC  
GGTTGCTATCGCTTCGCAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGC  
TGCTGGGTCTGCAGACGCGATGGATAACGTGTAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCA  
GTGTGAAAGGCCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCA  
TCGCACAAGCCCCCTGAACCATATATTGTTATCACTGGATTGGAAGTCACCGTTATCTTATTTTTTCA  
ACTTTTATATGT

SEQ ID NO: 834 ACAAAGATTGGTAGCTTTTATATTTTTTTTAAAAATGCTATACTAAGAGAAAA  
ACAAAAGACCACAACAATATTCCAAATTATAGGTTGAGAGAATGTGACTATGAAGAAAGTATTCT  
AACCAACTAAAAAAAATATTGAAACCACTTTTGATTGAAGCAAAATGAATAATGCTAGATTTAAA  
AACAGNGTGAAATCACACTTTGGTCTGTAAACATATTTAGCTTTGCTTTTCATTGAGATGTATACAT  
AAACTTATTTCCCGCGTACC

SEQ ID NO: 835 GGTACTATTTATTTTCTCAAGTGCTTCCATGGGGGAAAAAATAAAAGTCTAAT  
ATGCCAGAGAAATCATCATTGAACCAATAAGACACAGTAACATAATTCTAGTAACCTACTTCTCA  
ATGAACACACATCTGAGAAAAAACCGCCAGTATTTTATTCTCATGGAAAAACAGAAACAAACCCA  
CAAGTTGGAGTCACGGAGATAAAAATACAGATGAAATGGAAAACGGTCTGTTGTCTATGAACCTCTCA  
CTTTCAAATACCATTTTATATGGAAGTTACTTTACTGCGGGGCAAACAGAAGGCCATGCTGGAGTC  
TCTTACTTTTGGAAAATGGAGAATCAAAAATTTGCTAATCAACAAACAAAAAAGGAGGGGAAACT  
CCTTGGTAAAGCTCTACAAACATAATTATCATTTATATTTTACCAATAAAAGATACTAGGGTAGAA  
AAAACAGATGGGTAGAACTGGNGCCCAACCAAGTGAAAGCTTTNGGGCCTTCTNTAACTCCTA  
TCTGTTCTTTANAAAACCCAGTTTCTAAANAAAGATTCTNTGAATTC

SEQ ID NO: 836 ACGTGAAAGACGAATTTAGGAGACATAAGACCGTTGGTTCTGACGAGGCACA  
GCGTTTCTTGCAAGAATGGGAGGTGTATGCAACAGCGTTATTGCAACAGGCTAACGAAAACAGAC  
AAAATTCAACTGGAAAAGCATGTTTTGGCACCTTCTCCAGAAAGAAAACTTAATGACTTTCGTG  
ATGAACAAATTGGACAGTTGCAGGAGCTGATGCAAGAAGCCACAAAACCAATAGGCNATTTAGT  
ATTTCTGAGTCTATGAAACCAAAATTTTAGTCTATACAACAAAGCTTAATAAGACATGCAAAAATT  
TAGAACCCTACTTTAACTGTCATTGGTTTTTGAATATATTTAAGCTTTGAAAACACCTGTTATTA  
ATGAAATACTCTTTTATTTTGGATATTATGATTGCAGTATANTGGGATCAAGATCACTAGTGGACA  
ATTTGAAAAAACTATTGGAATAATAGCACTTGGTTTTAAAAATTCAGTTTTNGAACCTAANCAGCAA  
ATTTCTNGAATTTTGCCTGAAATGGTTTTNAAAATGGCTNTTCTCATC

SEQ ID NO: 837 ACGCGGGGGCTCTTCTAAGATGGCTGCCGCTACCGGTGCGGTGGCAGCCTCG  
GCCGCTCGGGTCAGGCGGAAGGTAAAAAGATCACCGATCTGCGGGTCATCGATCTGAAGTCCGA  
GCTGAAGCGGCGGAACCTANACATCACCGAGTCAAGACCGTGCTCATCTCCCGACTCAAGCAGG  
CTATTGAAGAGGAAGGAGGCGATCCAGATAATATTGAATTAAGTGTTCAACTGATACTCCAAAC  
AAGAAACCAACTAAAGGCAAAGGTTGTTACGAGTTATGAACATGTTTTAAAAATATATTTGGTTAT

ATAACTTGCCCTCTGCTNTTGTGCCTCTTGCCCTGTTATATAGTGCTACTCTNTGCTATTTTCCTCTGG  
NANTCTGCCTTTTTGGGTCAAGGAAGAAAAGTTCTGTTGCCANGCTTTGGGTGACTTGNACCCTTA  
AGATTNNGCTGAACTGANNTAAAAANTGAGACANTTCTTTATGTTTTCCGAAAAATTACAAAGTCA  
AGT

SEQ ID NO: 838 ACCAAGCACTGGGTAAGGCACTTTTGTGGAGCATTAGACAGTAACCCCTCAAG  
GAGCTAGAGAACCGGATGGGAGACATGAGCGTAATTAACCTCACTTGTTCCTCAGAGTTTCTTTTT  
GTTTTGATTTTCTTTTTCTGTGACTTATTTTCCTATTTTCTTTCCTCCATGTAATTTTCACTATGGCCC  
AACTAATATAAACACCTGGAAATTACAAGGAAAAAAATTTCTCCTCTAATAACTTTCCAAATTTG  
TGGAATATTTATTTGTAATAGCAGNTATCAGTTATGCTTATATAGCATTAAAAATTCTCCTCCTTTG  
ACTACACACAACCACAGTGTGGGTCTAATCATGGAGATATCAGNAATTTTANTAAACNTGAAN  
TTTGNAGGACATTTNTTTTGTACATGNTGCAAACCTGNTTTNAAATCTNANGTTCCAAANTAATT  
NTTTTTTTTTTTTTTTTGAATGGAGT

SEQ ID NO: 839 ACGCGGGGCTGTGACTTAAATCCATTTTCACTTAGAGAAATAGAAACACAAG  
GAAACCTTTGGACGCTTCATAACTGCTGGGAAAGGGGTATTATCAATTGTGTTTGAGAGTCAAACCT  
ATAAATTACTTCCCAAGGTTAGTTCTACCTATGCCCAGGAATGAACAAGGACAGCTTAAAGGTTA  
GAAGCAAGATGGAGTCATTTGGGTCTGATCTCTTTCCTGCTCATAATTTCTCAGTTACAATTTTTG  
TAAAGGTGGTTTCAAATGCTTTGCTGACCTCCCATAAACAAGGATGTGCCAATTGTAACCTCAGTT  
CTGCAATTCAGTCTGCTCCCAAACTAAAGTCCATTTGAATTGCATACTGATTGATAATGTCATTA  
CGAGTTTATCTTTCCATGTGCAGAGCAAAGACAAGAGGAGATCAGCCATTCTTCACCTACCCAGA  
GACGNTNGCATTGGAAAGGAGCCGTATCTTCTAATATGGTTTTTGCAAATGTAAAAATGTAAAAATCTA  
TGCNCGCTCCAGGANGGAAAAAANAAAAAANNTCCTCGNCGNGACCAC

SEQ ID NO: 840 ACTATTTATTTCTCAAGTGCTTCCATGGGGGAAAAAATAAAAGTCTAATATG  
CCAGAGAAATCATCATTGAACCAATAAGACACAGTAACATAATTCTAGTAACCTACTTCTCAATG  
AACACACATCTGAGAAAAAACCAGCAGTATTTTATCTCATGGAAAAACAGAACAAACCCACAA  
GTTGGAGTCACGGAGATAAAATACAGATGAAATGGAAAAACGGTCTGTTGTCATGAACTCTCACTT  
TCAAATACCATTTTATATGGAAGTTACTTTACTGCGGGGCAAAACAGAAAGGCCATGTGGAGTCTCT  
TACTTTTGGAAAAATGGAGAATCAAAAAATTTGCTAATCAACAAACAAAAAAGGAGGGGAAACTCCT  
TGGTAAAGCTCTACAAACATAATTATACATTTATATTTTACCAATAAAAGATAGCTAGGGTAGAAA  
AAACAGATGGTTAGAAACTGGTGCCAAACCAAAAGTGAAAGCTTTGGTGCCTTCTCTAAACTCCT  
ATCCTGTTTCTTTAGAAAACACCAGTTTTCTAAAGAAAGATTACTCTGAATTCACCAGGGTTCTAT  
CCCCCAATTCATCCCTCCCTTTCACCCCCAAGACACGAAAGGCCATGTAGTTTTTTGTCCGGCACC  
ACTTGGGAAGGGGCTGGCTCA

SEQ ID NO: 841 ACTGTCACAGAACTTTTACATACATTCTCAGTCCTAGTTGTGAAAGGCCTAAA  
GAGAAAGAACTCAATTTGCAGTCCAACACAAAGGGGGGAATTTCTAAAATAAATAATCCAACAG  
TTTTTTGCATTTTTTAAATTAATTTTTTCATTTTTTTTAAAAATAAAATAACCAAAAAAGTGTAAGTTA  
CAAAAAATGTCGTTGAAGAATAATNTATTAATACTGTGGAAAAAAGGAAAAAGACACGTCACA  
AACTTTTAAGATTAATATGAAGATCATAATTTAACATAAAGNGAATATATTCTATGGATTTGCCNT  
CCCGATAANTATGAACAATACT

SEQ ID NO: 842 ACTGGGACGATTCCGCGGAGCCGGGCAGAGGTTTTAGGGGAATGATTAACAA  
AGGCGTCCGAAGAAATCGTTGTTGGAAGGTGACCAAGGTGGAAGAGACGTTGCTTTGGCCCTGC  
AAGTAAGAAGAGAGAGGGAATAGCCTGAAGGAGTAACACTAAATTTAAATGACACTTTTTTACC  
AACCAGCGAAAGCAGATGTTCAAAGGGGATATTGGCCAGAGTCTGTACACTAAGATGAGAAATG  
TCCTTTCTTCTGAAAGGTGTCTGATGTGTAAAAATATGATATACTTTGTGCTGTTTCTCCTCCCTTCCCT  
TTTGCATATTATTCTGAAACAACATTAACTAGTTACTTTGCGTCATTGAAGGTATGCACTTCCCTC  
TATGTTAGGAGTGAATAAAATTAATAATAGATCCTTATAACAAAGAAAGGCAGATAGAATGATTA  
AAAATGACCAAAACATGTTAGAAACAGTCTCTCAGGTGTATGCAGATGGTAATTACAAAAATACT  
TTTTCAAAAATGATCTTCTGTGTCATGTTTCTGGGAACAAGTCAAGATGAATGAGTTTGATTTTTAA  
AGCAGAAGTAGTNTGTGTTGGTGTCTCATGAATACCACCAAAAAAAAAAAAAAAAAAAAAA  
NTCCTNGCCGNGACCACC

SEQ ID NO: 843 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTANAAAGGATGACTTTTATTTCCATCC  
TGAATGATTCACACCATTATTTAAACATCTGAAAAATCCTGAAATAATTTAACTGAAGGCACAG  
AACAAACCAAAATATTTAACTATCANAACTAAAAATGAGAAAAATCCAAATAGTTCTATAGTAACA  
ATAAATTATGAACAAGTTTCCGTCACCAAAATATCACTCTGACCAAAAAATGACTGTCTTTTGTCTATA  
AAAGCTACAGCTTAAGCTGATTCCAAGATTTCTATAAAAAATGAGAGTGAAGAAATTTCTTCTCA  
AAATACTCATTATGCCACCAGGTTCAATGTAAGTATTTGTATATAACAAAGTAGCAGTCAGGATA  
TTGTTGATGGATGGCTACTCCCCAAGAAATGACACATTCTTACAACTTTAAAAAATAGCAAAG

TTGGTTACAAAATTCTATTNNGGGAGCAGGGAAAAAACTTGTCCATGTCAATTAGAATATCAAAAG  
 CATTTCAAAAATCAAACCTGNTTCCAAAGCCAATCATCACACAACAGTAATATACTGTATGGGCCAT  
 TCCACTGNATCTTGTATATGTTGATCATAACAGNGTAACTAAGGGGNTGGAACAAGTCAGTANCA  
 ATGATTGTNGCTGTAT

SEQ ID NO: 844 ACTTTTTTTTTTTTTTTTTTTTTTTANATTGCTGCTGCTGCTAGGAGGAGGCC  
 TAGTAGTGGGGTGAGGCTTGGATTAGCGTTTANAAGGGCTATTTGTTGTGGGTCTCATGAGTTGGA  
 GTGTANGATAAATCATGCTNAGGCNAGGATGAAACCGATATCNCCGNTNCGGGNTGTATAGGATT  
 GGCTTGNAATGGCTTNGAGTNGGGNATCTGCTCTGNCCTTATCATTNATATTGATAACCTGAA  
 GTGATATTTTTCCTTACGNCCTTCCATGCAAATTACTCACTNNCCATAGCGAAGCACACGCNTTAA  
 TTGAAGTGAGAAGTAAANAAACNCTTNGTAACNGCTCNGATTNCTTATT

SEQ ID NO: 845 CGCGTCGAGGTACCCCTGGCAGAGCATTTCGAGATTAAAGAAGCATTGAGA  
 AAGAAGTTGGAATCATAAAAGCCAGCTTGAGAGAAAAGGAAGAAGAAAGCCAAAAACAAATGG  
 AAGAAGTCTCCAAACTTCAGTCGGAGGTTTCAATACTAAACAAGCATTAAAAAATTAGAGACT  
 AGAGAGGTAGTTGACTTGTCTAAATATAAAGCAACAAAAAGTGATTTGGAGACACAGATTTCTAG  
 CTTAAATGAAANATTGGCCAATCTGANTAGAAAGTTGANGAAGTATGTGAGGAAGTTTGCNTGCC  
 ATAAANGAAGGAAATTTCTGCANANNGATGNNANGGAATNACTGCATTTTACGATTGANCAANA  
 AATTAAGGATCAGAAGGAACNNTGTGTAAGTCCTTNCACCATCNCNGAGTTACTAAGAATACAAN  
 ATCTGNTTAAAC

SEQ ID NO: 846 ACGCGGGGAGTTGTCCTGCGCCGGTGTTCACGTCGCGCCTGAACCTGAG  
 CGCATAATGTNATGAGGAGATGGGAGCACTNNTGATTTCGCGGTATCAGGAATTTNAACCTANAGA  
 ACCGAGCGGAACGGGAANTCAGCATGATGAANCCCTCTGTCGCTCCCAGACACCCCTCTACCAAC  
 AGCCTTCTGCGAGAGCANATTAGTCGTGAGTGTCTATCCANANGTTAAAGGAGANATNGCTNGTA  
 AANATGAATAGCTGCTTGTGCGNGTCTAAAAGATGTGTATGTNGATTCCANAAATCCTGNGTNTTAC  
 TTG

SEQ ID NO: 847 ACCCTGACCCATGAACACCTGGCCATGACCTTTGACTGCTGTTACTGTCCACC  
 TCCCCGTGCCAGGAAGCTATTTCCAAAGAACCTATCGTGATGAAAAATTTATATTGGATTGAGAA  
 AAACGCCTATTCCCATAAAGAAAACCTTCAATTAAATCAGGAGACAGAAGCCATAAAGGAAGAA  
 CTGTTGTATTTTAAAGCTAATGGTGGAGGGGCTTTGGTGGAAAACACAACCACTGGGATTAGCCG  
 AGACACACAGACGTTGAAGAGGCTTGCAGAAGAGACTGGCGTCCATATCATATCTGGAGCCGGGT  
 TTTATGTGGATGCAACTCACTCCTCAGAGACCAGGGCCATGTCAGTGGAGCAGCTTCCGATGTCCT  
 TATGAATGAAATTCTCNTGGAGCTGATGGANCCAGTNTCAAGTGTGGCATNTTGGAGAAATTGGT  
 TGTCTGGGCTTTGATGAGAGTGAAANAA

SEQ ID NO: 848 TCGCGGCGAGGTACTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTGCTTTTATAGG  
 TATCTATCTAATAAAAGTTTATTTGTGTATGTGCAATGCATAACTCTATCTTAGATATGAATCCTAA  
 CAGGATGAAAAATCTTTCTTGCAACTACTTTATGCTTATGAAAGGTGTGAAGTTGCAATGTCCTCC  
 TGTCTTAAACCCAAGTTGACAGNGCCCTCTCAAAACTTTTCATAAATAATGACCTAATTNCATTTA  
 AAAAAATGGTTTCANCAATATGAAAATAGAAAGNCCCGTATTTGCCATTTGTAATATGAGAAAAA  
 AAAGATGATACATTCTNTACAAAAAAAAGTGGGTTTAGANAACAGTTCTGGTAGTATTTCCATG  
 GTAAAGAATCNAAAGATCTAATGACNGCCCCCTTGCTCNGGGAAGGACAGGGAATTCATGTGTTT  
 TTTTCCGATNGCTTAGNAACTCCANGGGATTTCAATTGGGGGNCAATACTGGCCNNGNNTCAAA  
 TTTAATTTNTTNTNGGNCCTTATAAAATCAAAACTTTTTACCCTNCAATTGGGGGGNTTGGGTA  
 GGCCNTTGGCCAGTCTAT

SEQ ID NO: 849 CCGCGGCGAGGTNCCCTTNNNGAATTTGAAGTGAANGATCCTGAGCTGGAGG  
 CCCAGGGAGATGACATGGTTTGTGATGATCCGGAGGCTGGGGAGATGACATCAGAAAACCTGGTC  
 CAAACTGCTCCAAAAAAGAAGAAAAATAANGGGAAGGAGGTTGGAGCCTTCTCAGAGCACTG  
 CTGCCANGGTGCCAAAAAAGCGAAGACATNGATTCCCTGAANTTCATGATNCNAANTCNGATGTG  
 TTTGGNCCATTTATNCCGNGATCCNACATCATCCCTATAATGATCTGCCCGACTGGAGCGTGCT  
 CTTCANGATCCAAATGTGGCTGNCGTTTNTGNAACCAATTGAGGGGTGATCAAGCNTTTTTGT  
 CNGATCCANGTTACCTATGGGANTGCNGGAGCTTNCACANNCNCCAANTCTTTTATGCTGATNAA  
 ATNCAANCTGATTGCCANACNTGGAAATGNTGCTGTNATTTGNAAAATGCNGACCT

SEQ ID NO: 850 ACTGTGAAACCACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCA  
 CCACCATGAAACCTACAGCGGCATCTAATACAACAACACAGGGATGGTCTCAACAAATATGACT  
 TCTACCACCTTAAAGTCTACACCCAAAAACAAGTGTTCACAGAACACATCTCAGATATCAAC  
 ATCCACAATGACCGTAACCCACAATAGTTTCAAGTACATCTGCTGCTTCATCAGTAACAATCAAC  
 AACTATGCATTCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGNTGGTGGTATTG

NATTAACGCTGGGANTTTTATCTNTNTTTACATTGGATGCCAAATGGNTTNTCTCAAGAAGAGCTT  
TCGGGGTTTCNAACCATAAATNAACATGATGCCTTTTTTAANGAATCCATGGACCAAGGATGGAT

SEQ ID NO: 851 CCGCGGCGAGGTTNCCCTTNGNGAATATGAGGAGTATATTACTAAACTTTTCA  
ACTACCACAAAGTTCTTCTATGAATACAGGAGTGGAGGCTGGAGAGACTGCCTGTAAACTAGCT  
CGTAAGTGGGGCTATACCGTGAAAAATAAAGGGAAAAAAGGGTTGGAGCCTTCTCAGAGCACTGC  
TGCCANGGTGCCCAAAAAAGCTGAAGACATANGATTCTGAANTTCANGATCCNAANTCCGATGG  
TTNGGNCCATTTATNCCGGGATCCNACATTCAATCCCTATAATGATCTGCCCCGACTGGAGCGTGC  
TCTTCANGATCCAAATGNGGCTGNGTTTNTGNGNAAACCAATTCAAGGGTGAATCAAGCGTTTTTTG  
TTCNGATCANGTTACCTATGGGANTGCNGGAGCTT

SEQ ID NO: 852 CCGCGGCGAGGTTNCCCTTNNNGAATTTGAAGTGAANGATCCTGAGCTGGAGG  
CCCAGGGAGATGACATGGTTTGTGATGATCCGGAGGCTGGGGAGATGACATCAGAAAACCTGGTC  
CAAACCTGCTCCAAAAAAGAGAAAAATAANGGGAAAAAAGGGTTGGAGCCTTCTCAGAGCACTG  
CTGCCANGGTGCCCAAAAAAGCGAAGACATNGATTCTGAANTTCATGATNCNAANTCNGATGTG  
TTTGGNCCATTTATNCCGNGATCCNACATCAATCCCTATAATGATCTGCCCCGACTGGAGCGTGC  
CTTCANGATCCAAATGTGGCTGNCGTTTNTGNGNAAACCAATTCAAGGGTGAATCAAGCCTTTTTGTT  
CNGATCCANGTTACCTATGGGANTGCNGGAGCTTNCACANNCCNAANTCTTTTATGCTGATNAA  
ATNCAANCTGATTGCCANACNTGGAATGNTGCTGTNATTTGNAAAATGCNGACCTNNNATAGNC  
CTCTNGAAAGGCCTTCNGGGGGCTTNTCTTGGTTTGA

SEQ ID NO: 853 ACGCGGGAGGATTGTTCCACTAAAATTTATTTTTCAAAAAATTTACTTCACAT  
TATTCTATGTAAGTGATGACTTGTGAGTGTCCAGGTGTATCTTAGCTAAAACCTAGAGAATGCCCT  
AACTTAGATGGTTTTTTGAAGCCTATACAATTGGTATTGTTTGACCCTTAAGCTTTTACATCTCTTAG  
CATGGAGGACGAAGAAAGCTGT

SEQ ID NO: 854 ACTATTTATTTCTCAAGTGCTTCCATGGGGGAAAAAATAAAAGTCTAATATG  
CCAGAGAAATCATCATTGAACCAATAAGACACAGTAACATAATTCTAGTAACCTACTTCTCAATG  
AACACACATCTGAGAAAAAACCGCCAGTATTTATTTCTCATGGAAAAACAGAACAACCCACAA  
GTTGGAGTCACGGAGATAAAATACAGATGAAATGGAAAAACGGTCTGTGTGTCATGAACCTCTCACTT  
TCAAATACCATTTTATATGGAAGTTCTTTCTGCGGGGCAACAGAAAGGCCATGCTGGAGTCTCTTA  
CTTTTGGAAAAATGGAGAATCAAAAAATTTGCTAATCAACAAACAAAAAAGGAGGGGAAACTCCTTG  
GTAAAGCTCTACAAACATAATTATACATTTATATTTTCCAATAAAAGATAGCTANGGGTAGAAAA  
AACAGATGGTTAGAACCTGGTGCCAAACCAAAAGTGAAAGCTTTGNGCCTTCTCTAACCTCCTATCC  
TGTTTNTTTANAAACCCANTTTTTTAAANAAAGATTNCTNTGAATCCCCAGGGTTTTNTNCCCA  
ATTNTTCCCTCCTTTCCCCCCANAAACACNAAAGGGCCTTNTNTTTTTTTGCCGGNNCCCTTGGGAA  
GGCCTGGGCCTCANGGGCC

SEQ ID NO: 855 ACTACTGTTAATATCTCTAAGAACAACACATTGAACATCCTTCCAGAAAG  
TCTTTGAGGGAGGACCTATACCCATAATAGAATTATGGCACTCATTTCTGACAGTGATCAAGAAAT  
CAGTTATTTCTTACTGTTGGAAGGACATTGTAAAGTATGTGGTTATATGCAGTGAAACTGCAGAA  
AATACTCCTGGTTGAGGAGTTTTCACTTTACTACAGTGATATAAAAACCAGCAGTTTTTACACTAA  
ATTTTTAAAGAAATATTAGACAAAAATATAGAATTAACCTTTGGTTCCAAAATGGGAAAGGT  
TCCACGATACATAAATCATTTCTCATTTGCTTTAAAAAATTTAAAAGTGTAATAATTATGAGAGAC  
TTTATTCGTTAACAATGGGGGTAAAGAGCTATATACATGAAAATGAGTCTTATAAAATTAAGTGA  
AGTGCAAAATAAAAGCACTGCTACTATAAGACATTCTGGAATGGTTGTTAATAAGGGGATTATCCA  
TTTGATCTATAGCAATGTGATTTTATTTTTAAAAAGAAAAGCAGGGGGTTTTCTTTTTTTGGTTGGT  
TTNCTTTTGCTTAANCCCTTCATCAATTGCTTTATCTGGATCTGCGAAGTAATCTTCAAATCTCTTTT  
GTTCTTTTAAAAAT

SEQ ID NO: 856 ACCCATCCCAACTCTCAAATCGTTTGGTTTTTTTTTATCTTGATTGAGATCCTC  
TTCTCACTATGCTAGTGGTGGAGATATTGACAAAATCCTATTTCTTTCAAAGAGGAACCTTTTACACA  
CCGAAAAAAGAGCATGGAATTATTTTATATTGTTATAAAAAATCCAGATGCAAAATTTTTTAAATGC  
CAATTATTAGAGCTTCTGGGGAAAAAGTATAGTTTACGGAAATAAACTATGTTCTTTTCAGGGTTG  
GGTGGATAGGTGGCTGCTAGGGTGCTGGCTCCTGGCGGCTTTGCCATCCATGANGCAAGGGCTGG  
GAACACAGTGCTTTGCCTATGGTAGATCCATGTGAATGCAGGAAGCCAGCTCTTCAGTCTTGGAGA  
TGATTTCTGCTACAATTCTGTANAAAGATTAAGGATGGCAGAGTAAAAGGTACCAAGAATGCCAG  
GATGTTTTTCTTGGCCGTAGGANGTCCAAATTACTTTNCTTTTTGATGAAAGAGTTTGGAAAACTG  
TCCATCTCTCTGGCTTGAAAAATCTCTGCCATTTTAACATCCTGNGAAATAGCAATTATTATCATCT  
GTATTTAGTTTAACATTACCCACACCTTANAAATAATAGGTNAAAATCGCTTGCTACTCTTCCAA  
AGATGATCAAGTCAT



SEQ ID NO: 857 ACTGGTGCACCTCCTACATATCAAGGAAAAGCAAAACACAGAATAATTTAAT  
ATGCTGAATAAAATTTGTTTACACCAGATACTATACATCACTGATGGTTGTCCAGTATGAATTT  
TAAGGGTATTATGTTAGATTCTGCAAAATATATTCCTATTATTCACAAAGTGAGGAGTCAAAGTCCA  
ACTATTCAAATGGCCATAAAACAAAATGTTGCAGAAAGGTATAGACCATTAAAAATAAAAAAGTCA  
GGAGTGGGGCAGCTGACCCCTTAGGAGCCTCANGAATTCCTTTTAAATGCAAGATAGATGGCAAGA  
GCTGGCTTTTTGGTTAAGTCACCAGTCGGAAACCCATCAGGGAAAAGTTTTCAGNTTCAACTGGT  
AAGGCAATGACTTCTTNAAC

SEQ ID NO: 858 ACTTGAAGGAGAACAGTTTACATCGGGCGTTAGCCACCTTGCAGGAGGAGAC  
TACTGTGTCTCTGAATACTGTGGACAGCATTGAGAGTTTTGTGGCTGACATTAACAGTGGCCATTG  
GGATACTGTGTTGCAGGCTATACAGTCTCTGAAATTGCCAGACAAAACCCCTCATTGACCTCTATGA  
ACAGGTTGTTCTGGAATTGATAGAGCTCCGTGAATTGGGTGCTGCCAGGTCACCTTTTGAGACAGAC  
TGATCCCATGATCATGTTAAAAACAAACACAGCCAGAGCTATATATTCATCTGGAGAACCCTTTTGGC  
CAGGCTTACTTTGATCCTCGTGAGGCATACCCAGATGGAAGTAGCAAAAGAAAAGAGAAAGAGCAG  
CAATTGCCCAGGCCCTTAGCTGGCGAAGTCAGTGTGGTGCCTCCATCTCGTCTCATGGCATTGCTGG  
GACAGGCACTGAAGTGGCAGCAGCATCAGGGATTGCTTCCTCCTGGTATGACCATAGATTTGTTTC  
GAGGCAAGGCAGCTTGTCAAAGATGTGGAAGAAGAAAAGTTTCCTACACAAGTACAGGAGCATAT  
TAAGTTTGGGTGAGAAATCACATTGTGGAGTGTGCTCGATTTCTCCANATGGTCAGTATTTTGGTC  
ACTGGGTCTGTTGATGGA

SEQ ID NO: 859 ACCTGCCTTGAAATTTAAATGTCTAAGGAAAATGGGAGATGATTAAAGAGTTG  
GTGTGGCCTAATTCACACCAAAATGTATGCATTACATCCTGCTCCTTTCTAGTTGACAGGAAAGAA  
AGCTGCTGTGGGGAAAGGAGGGATAAATACTGAAGGGATTTACTAAACAAATGTCCATCACAGAN  
TTTCCTTTTTTTTTTTTNGAGANATATCTGGCTCTCGTCACCCANGCTGGAATGAANTGGTATGAT  
CTCAGTTGATGNGCANCTCCACCNNCTANGTNCANGCNATTCTNATGCCTCACCTTTNAACNGNT  
GGAACATANGCCCATGCTACCNTGCCAGGCTANTTTTTATATATATANTAAAGNCGGGTNGTGGC  
NANTGNTNGCCAGANATGNTTNGAACTCNTGGCCTAAGATNAATCTGCCCACCGTNACCTCCCTA  
AGTGCTG

SEQ ID NO: 860 ACTCTAAGTCAGGAAAAATTAAGACGACCAATAGTTGCCTGTGAACCTGGCA  
GACTTTATAACAAAGATGCCGTCAATTGAATTTCTCTTTGGACAAATCTTGCAAAAAAGCTCTTTG  
GGAAAGCACATTCTCACATTTAAAGCNNTTAAAAATGGTGACAAACCTGANNCTTTCTGNNTAAT  
CCTGCCTGGGAAGGGGATAAAGGAACCACTAAAGGTGACAAGCACAATGACCTCCACCGGGCCC  
NTTNATTTTGCCCGTTGTNGGCCNTGGAAATTNACCGGCCAACNCAGGTCTTGTTCCTTCG

SEQ ID NO: 861 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCCTTTAAAAGAATTTATTAAGCCTGTT  
ATACCACACAGTNTGTTTATACACTGACATACANCTCCNTATTAAGATAAAGCAAAAGACAAAAA  
AGTTTNTNTNTTAGAAACAAGATNCCNCCANTTATTGNTNNAANCATTATTGCACTTTANCTT  
TTTTANTNGACAAAGCATTNAAAAAACATTTGCAAACTAGTTTTAACAACAANTANCACCTGTA  
NCCAAANACNTNACNGCCCNAAATGGTTTATTANGTTGNATTTTAAACAACCTTTNCTTTNTTNACC  
GNAACCGGGGGAACCTGNA

SEQ ID NO: 862 ACACAAATGCATGAGTATGTTTATACAGTGTTAGACTGATGTGAATTTGCATT  
TGTTACATTACATTGCCAGCGCATATCATTTAGCAAGTTGGCATTAAACATTTATGCTTTAATTAAT  
GCCAGTATACCTATGTGTGCAGCAGTAAAAAATTAGTGAGAAAAAGCAACTTTTGTCACTCTTAG  
GAAATATTTTGTCTTATTAAGTGTTCTTGGCACATGTATATTACTAAAGTANATAATTCCAATGAG  
AAATACTCCAGATTATTGGTATAAAATTAATTTACAATGTCCCTGATATTGACTACTCTTAAAAAA  
ACCAAACAAAACCTCGTATCTGATGTAACTTTGCCAATATTTTAAAAGCCAAAATATTCTNTGGACA  
CAAATTTGTTTGTNTCAAGGACAGGTTACCTTGCTGTTAAACCTTCCNAACAGAAATATACTATCT  
ATCTTTGGGTTTGGTTTTTGGTTTTTGGTTGGTTGGAATAAAAGGAATTCCTCTTGTCCTCCAAAGCT  
NGATGTATTGGCGCATCTTAANTNACTGGNANCTCANCTCCNGGGTCAAGGATTCTCCGCTAACTT  
CTGAGTACTNGAATTAANGNCCNCCCCCNCCCCCGGTAATTTTGGGTTTTANCAAAACAGGGTTC  
CCNCCNTGGCCAGG

SEQ ID NO: 863 ACAAAGGCTGCTTAAGGCAGTGCAGCCCTTCTCAAAGTCAGCATGTCAATG  
AGAGACTGCTTGATACTTGTCTTTCGAAAGCTATGTTTGCCAACCAGCTTGATGCCCCGAAAATCT  
GCAGTTGCTGGGTTTTTGTCTCCTGAAGAACTTTAAAGTTTTAGGCAGCCTGTATCCTCTCAGT  
GGCAGTCAGTCTCTCAGGTCAAAGTCAGGTTTATGTGGATGTCACAGCCATTACAATTCTGTGCGCA  
TGAAACCTTTTGCTTGAATCATGGATAGTTGAGGAGATCTTAAACCCANCAGCTGATGTCCACTA  
TGCTTATGAAGGGTTTATATGTCTTCNAAGAACTCTAACTGCTATTTCAAGCATGCCAACTNTTGT  
NTTACAGTTAAACAATTCTTGACCCAAACCGACTGNTGCCTCTNTGAAAATAAAAACCTGGANTCT  
GACCAAGGANATAAANCTCTTACAAAACCACTGGATATCTGCTGNGTGNATTCAACATTTGGTTGG



CCTGGTATAAAAATACAGCNTACCCTACCCNNGGAAAGGAGGAAAAGGAGANGAAAAGCTTTTT  
CCAAAACCTAATTATTTTTGGAGNCNTTCTAATNNAATGATTAAAA

SEQ ID NO: 864 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAATTAATCCAAATTTTATTA  
AGGATTTTCAGGTTACATACTTCAAATTTCTAGAATGGAATGGAATCATTTTGGAACTGGAAAAATG  
GCATAAACACTGACGTCCCTTAAAACTTCAATTTTATAAAAAAAATTTCTTCTGCAAACCATCC  
CTTTATGTAACAAGACTAGGTATTATCTACACCTTCACCTTGGCAATAGCTATTTCTAAANAATG  
AAAAANATGATTTTGCTACTTCAGTTCATTNAAAAATGGGATTCTATCTTTGAAGTTCAAAAAAGC  
CTCATTTTCGATGAACCATNGTTAAAAAAGCNCATANTGNTAATCAAAGCNAAGGGAATCT  
TTTTGAAAATNAANAAGCTNTTATTATTGGGTTAANGNTTCCAGNAGTNAGTTAAAAATCT  
NAGGAGANGAATGGATNGCCCTTCAAACAATNTNTTCACATTTCAAGACCTTAACACTACTTTTCA  
AAAAGGGAGTTTCTTCNNTAAAAACAANTCNAATNTATTTTTTAAAAACNNGGAATNTAAAC  
TNCCTTTATTANAATTCAANTAATNCCCCTCCTTTGNAAAAAAGTCNACCCNAANTAACANNNT  
TTNCCCCGGGTNTTG

SEQ ID NO: 865 ACCATGAAGAAGAATAAATGAGGGTAAGGGGCTAGTGTGATAGGGAGAGGG  
GTGGGATGCCATCATATTTAGGGGTGGTTGGGAAGTGTCTACTGCTTTAGCATTGTGTCTTCAA  
TTTCTCTCCTTTGGTTATATACCTTGCCTATTCCGCACATTGATAAAGTTTCTTTCTACAGAAGTTC  
TGATATTGAATTAAGGAATGGGGTCACTACTTAAGACTTTATCATTTTCAGCTACACATAAAGGTT  
TCTCTCCCTATGGATTTTGCTAATGGTTGAGTGATACCTAAAGGCCTTGTTCATTTTACACTT  
GGGGTTTCTCTCCGNTTGAATTCTCTCATGGGTACCAAACTCTTTTAAAAGGACTTTCACANGC  
TAAATGTGTATGGCTCTTTCCCAATATGAATTTTTTGTATGTAAGTACTTAACTTAATTA  
AACCTNTTGTCTNCTGCATTATAGGGGTTGGGCCNCTGANAATAATTGATGCCTGATTNAGGGC  
CAATCATNACCAANACTTTCTTTCTNTTTTATATAAAAAATGAAGNGAAGNCCNGGGCTTAA  
ACCCNNAATAACTTTACTTTCTACCTTTACANGCTTGCTGGGAATTNCNTGGGGTNNATAAAG  
NGAACCTGG

SEQ ID NO: 866 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNAATAAGGNCTCACTCTGTCATCCAGG  
CTTCAGTGCAGTGTGATCATAGTTCTGTAACCTAAACTCCTGGCCTCAAGCAATCCTCCTGCC  
TCAGCCTCCCAAAGCGCTGGTATTACAGATGTGAGCCACCAAGCCCAGCCTAAACAAGCATTTCT  
CTATTAAACCTTTTTCANAAAGACTGNATGAATTAGCCCAAAAGTGNCTTTTTTTTTTTTTTGA  
GAAAGTCTCCCTCTGGTTCCAGCTINGAATGGAAATGGCGNNAATTTCCGNTTNNTGAAATCTC  
AACCTNCTAAGGTCAAACAATTTCTCCTGGCTNAACCTCCGANTAACTGGGACTAAAGGGGC  
CCACACNANACCCCGGNTTAATTTGTATTTTANNAAAAAAAGGGTTTACATATTTGGCCAGC  
TANTCTNAACCTCCGGACTTTGNGAACCNCTGGCCTAAACCTNCCAAAGTNGTGGGATTANA  
GGGGNGAGCCACTTGGNCCANCCNAAAGTCACTTTTGTAAAAAAATTTTACNCTAAAAA  
AAGATAANAATANTCTTGCAATCANAAATTTTTNNAGAAATCAACAANANTGTTAAAAANAAGT  
GACCTTTGTTTTAAAAAN

SEQ ID NO: 867 ACATAGGGTCTGTACACCAGTTTTAGGATAAAGAACTGGAAGAATTCCT  
CAGGGATCAGTCCATAACGAACTTTTCTCCGTATTCAGGAAGAGGTGGGTACGCGGGGGAGGCA  
TTGAGGCACAGCGCAGGGGCTTCTGCTGNNGGGGCAGGCGGAGCTTGAGAAACCGNAGATAAG  
TTTTTTCTCTTTGAAANATAGAGANTAATCCACTCTTAAAAAATATAGCCATAAGGTACTTAAG  
ATTTGCTTAACCGTTAATTTTACCGTATTTAATAGCTTAAGAATTTAAAAAAATTTTAAAGACT  
TAAAAAAGTACNTGAAGAAAGGAAANATTAAAGGGTCTTAAACATGACGGAGGTTGAGAATA  
AACCTCTTCTGGANTAAAAAATGTTTTNAAAGAAATTTGAAAGAAGGCTNCNGACCCGATTAA  
TTCCAATAGAAGGCCATGCTTTAAATAAAATGAAGGNGCTTAAACAGCTTAAGTTAATTTAAAA  
GTTGGAGGTGATAAAAAAAAAAAAAAAAAAAAAAGTNTCTCGNCCGACCCCNCTANGGGCA  
ATTTCAACCCACTGGCNGCCGGTTCTTAGGGGATCCGACCTCGNCCCAANCTTTGGGGTATTCTG  
GGCTACCTGGTTCTTGNGNGAA

SEQ ID NO: 868 ACCTGTCCCATTCCTAAAAGGATTTGTGGGCAATGCTGGCACTTGGTGGCCAG  
GAGAATCTTCTGACCCCACTCTCCCTCCTCTTCAGTCTGAAGACCCCAAGAACCCAGTTAGGATC  
CCCTGGCCAGAGGTCTCTGTGACTGCCTCTGGACTCAACACGTGCAGCAACTTGGGAAGAATTTGA  
GCCAGTCTCAAAAACTTTNNCCCCCAGAATGAGAACCAGTGACCCCAAGCNAGGAAGGGCTGG  
GGAATCTGGAANGGAAGANAGGGGGTCCAANGGACCCTGTTGGCTTAAGCCTTGATNACC  
AG

SEQ ID NO: 869 ACCCACTGCTATTGCCTAAGGGTGTAGTCCTGAAACTGAAGCCAGTTGCCGA  
CCGTTTCCCCAAGAGGCTTGGAGACAGAAGCGTTCATCAGTCTGAAACCCCTCCTTATCCAACC  
CAGCCCCCTCTCCAGCCAGCTTCAACCTGGGAAAAACACCAGCCCAATCAACTCATTCAGAAG  
CCCCCTCCGAGCAAAATGGTGTCTCCGATTCTCACCAATACAGCCAGCCACTGTTTTACAGACAG

TTCCAGGTGTCCCTCCACTGGGGGTCAAGTGGGAGGTGAGAGTTTTGAGTCTCCTGCAGCACTGNC  
TGCTATGCCCCCTGAGGCCAGGACAAAGCTTCCCTCTGTTGAGTCCCAGACTTTGCTCTCTTNTGC  
CCCTGTGCCCAAGGTNATGATGCCCTNCCCTGCCTCTTTCATGTTTNGAAAGCCATATNTGAAGAC  
NGAGACCCTCNAAAAGAAAGGGAAGCCANGGCTTCTGTTGTTTANAACCTGCCCTGTTATCAC  
CCTNNATCTGTATCNTTACTGGTCCCGCTACCACTGGTGAANAATTGTGACCCTTGGNCG

SEQ ID NO: 870 ACAGGCTTAAATCTATGTCAATTTACACTCACTGAATCATCAACCTNATCACCA  
CCTGTTCCCTGGGTGTAGCGGGTATNATCTGCTTCCCTCATCATCATCATTTGACCAGGTCAGGACGA  
AATTCNAACACTTTACGACCACATGATCACTAGTGCTTTCCCTGCTTTGAAGTCAGCTTTCCTTCTT  
TCCATATNTTGNCAAGTTTATCAATCTTTTCTTGNCTTTTCTTTTNTTCCAGGCAANAAAAGATT  
CTAGAGTGATTTTGGTAACATNTTGGACCTAGGGCANANCGCTCTCTCTCAAATTAATCTTCTAA  
TGAAAATTNCTNNTTNTCTCINCTTTCTTTTATCTANTTTCAACACATAATNCANGAANGAA  
GTGCATGACGAATACATGCNTAATATCACCCCTCTTTCANGGGCATACCCTAAACACNNCATTANT  
TGGATGTTTTTCAATCAGNTATCAGNAAATGCTTGGCCCANTTTTANAGTTTTTGGTNTTTTTCTT  
TTCCNNCTAACCCGTNNTNTTNGGNTAACTANCTATTTCCNCTTTTTCTNNATCCNCG

SEQ ID NO: 871 ACTTACCTTCACCTGAAGAGCGTAACTATGAATTTCTGCAGTGGTATAAGGAT  
GACTTATGTGCATTGGCATTAAAGTTCTCCATGACAAGCAGCGAGGACCACTGGTTTTATGCGC  
ATTTACTCAGGCACTATAAAACCCCAAGTTGGCCATTTCATAATATTAATGGAAACTGCACGGAGAG  
AATAAGTCGTCTGCTTTTGCCGTTTGTGACCAACATGTAGAAATCCCTTCACTGCTGCTGGTAA  
CATTGCTTTGACTGTTGGCTTAAACATACTGCCACTGGAGACCACTTGTCTCATCCAAGTCCAG  
TGCATTAGCTGCAGCTCGTAGAGCCGAACGGGAGGGAGAAAAGAAGCACAGACAAAACAATGAA  
GCAGAGAGACTTTTATTGGCTGGAGTGGAGATTCCAGAACCTGTTTTCTTCTGTACC

SEQ ID NO: 872 ACTTATGAACCTTATGTTGCTGTTTACTTCCCTTTTCTGATTTTTTAAGTTCATC  
ATACTTTGACGAAATATGATGGGCATACATTGGCCCTAGACAGGAAAAATGCTGCATCATGGGGTG  
TATGAAATTCATTACAGAGTACGCTGGGCAGAAGCAGAATGTCAGAGACAACAATTACCTGTGAC  
TTTTGGGAATAAAACAAAAGTTCTAGGAAAAGCACTTTCCTTAATCCGGTTCCCACTGATGACAAT  
TGAGGAATTTGCAGCAGGTCCTGCTCAATCTGGAATTTTGTGAGATCGTGAAAGTGGTAAACCTCTT  
TCTTCATTTTACTGTCAACCCTAAACCCCGAGTTGAATACATTGACCCGACCAAGATGCTGTCTCA  
GGGGAAGGAATGCTGCATCAATAGATTCCAGCAAGTAGAAAGCCCGCTGGGGTTACAGTGGGA  
CGAAGTGATCGAATCAGGCAATATAAATTATCAAAATTGACTGAAGAAGAAAAGCCTGAATAG  
ACCAATTACCATGGAAGGAATTAAGAGNTATCAAGAACTGTCCTTCCAANGNGGGTTTACAGA  
TTCACAGTTAATANAAGGACTCTATAGTTGGATTGGCTGNATGGACTATCATGGCCTACGATATC  
AAGGGATATCNGATCATTGAAT

SEQ ID NO: 873 ACGCGGGGGGACAACCTGGCCATCCAGACCCGGGGTGGCCCAGAAAAGCAT  
GAAGTAACTGGCTGGGTGCTGGTATCTCCTCTAAGTAAGGAAGATGCTGGAGAATATGAGTGCCA  
TGCATCCAATTCCCAAGGACAGGCTTCAGCATCAGCAAAAATTACAGTGGTTGATGCCTTACATGA  
AATACCAGTGAAAAAAGGTGAAGGTGCCGAGCTATAAACCTCCAGAATATTATTAGTCTGCATGG  
TAAAAGTAGTCATGGATAACTACATTACCTGTTCTTGCCTAATAAGTTTCTTTTAAATCCAATCCAC  
TAACACTTTAGTTATATTCACTGGTTTTACACAGAGAAATACAAAATAAAGATCACACATCAAGAC  
TATCTACAAAAATTTATTATATTTACAGAAAGAAAGCATGCATATCATTAAACAAAATAAATCT  
TTTTATCCCAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG  
TCACACACTGGCNGGNGCTTCTAAGTGATCCGANCTCGGGACCCANCTTGGNGTAATCATGGGCA  
TAGCTGNTTNTCTGNGNGAAAAATGNTATTCCGNTTACAATTCNCACAACATTCCANCCCGGAAGNT  
AAAGNTNAAACCTGGGGNGCTAATGAG

SEQ ID NO: 874 ACGCGGGGGGTTCTTGGCTTTGACAGCTTCAAAGAATGGACAGTGATAAGTT  
AAAAGAAATTTGTATATTGTCAAGGAAAGGGTCTTAAATCCGAGTCAAGTCCCTTCCCTTGGGGTA  
AAAAATGTATTCTTAAAGCATTCTGATGTTAAAAAGAAAACCTTAAAGTTATCTAACCAAAACAGAC  
GCAAGATTTTGTCTGACACTACTTGGCAATCAAAGTGATCATAAATTAGGTTATCAGTTTT  
CAGAAAGTTGCTTTGTGAGAAAATTTGTTAGATATATTCTCCCAAGCATGCTTTTGTGGAAGGT  
TTTCAGCCATTGCCACTGAATCAGATGTTAAAAATGAAGGGAAAATTGAGTGTGCACACACACAA  
CTGTTGTACC

SEQ ID NO: 875 ACTGTCCATATCTTTTGTATTTACTTCAAAGGATTCTGGATCAGCAGTATAAA  
TAAGATTCTCAGCATCTGCTTTACAAATGGTGTTAGCTACATGTCGACACAGCATCTTTAGCCAGT  
TTTCTTTTGGAAAGTTCATCTGATGTCTATCTGGAACTGAGTAGCACATTTGCCTGCTCTGTTGGTGG  
CCTCACAAGCAAGGCAAAAGCATTATGGCAATCTTCTGCTCTCTTATGTCCAATACCTTCTTAAT  
TGAAAAAGAGGCATTAGGTGAATATGCTTAAGAGAAAGCTGGGGTGGGGTGGGGTGGGGTGGGGT  
CCTAAAAGTGCCAATAACCTTGTGCCGTTTTCTTGTCTATCTTAGGCAATCATTGAAGAGGAAGAG

AGTTACTTGTTCTCCTCTGTCACAGGGGGTGCTCACCTAGAGAAATTGTTTCAACTCGCTGTACC

SEQ ID NO: 876 ACGCTGGATAGCCTCCAGGCCAGAAAGAGAGAGTAGCGCGAGCACAGCTAA  
GGCCACGGAGCGAGACATCTCGGCCGAATGCTGTCAGCTTCAGGAATCCCCGCGTACAGATGAA  
TCGGAATTTGTCTAATGAGGAGTTAACAAAATCAAAGCCATCTGCTCCACCCAATGAAAAAGGAA  
CCAGTGATTTACTTGCTTGGGACCCCTATTTGGACCATCTCTTGATTCATCTTCTTCATCTTCACTA  
ACTTCATCATCATCAGCCAGGCCACAACCTCTTTCTGTAGGCACCATTGTCCACCTCCGAGG  
CCTGCTTCCAGACCAAAGCTTACTTCAGGCAAACCTCAGTGGGATTAATGAAATACCCAGGCCATTC  
AGCCACCTGTAACTTCCAACACCAGCCACCTCCTGCTGCACCATTAGCCCGGGCAGAAAGTTCT  
TCTTCTATCTCATCATCTGCTTCATTGAGTGCTGCCAATACTCCAACAGTAGGTGTGTCACGGGGTC  
CCAGCCCTGTCAGCCCTTTGGAAATCAGGATACCTTACCTGTGGCAGTGCCCTTACAGAATTGTTAT  
GCCTACTTTAAAGGAGCAGATCCCCAGTGTATTGNGAANATACTGGNGATATGACAATGTCNTTTC  
AAGTGGAATTATTAT

SEQ ID NO: 877 ACGCGGGGAGTGTTTTCTGGGATGGGAACCACGCCGCTTCCCAGTCTCTGTGC  
GAGGCGTGAAGCGCGGACCTTTCAACAAGGGCTTTATTAATTCTCACGCTGCGGCCCTGGAAAGC  
GATGGAGGTGGCGGCTAATTACTCCCTACGGGTGAAGAGACCTCTGTTGGATCCCCGCTTCGAGG  
GTTACAAGCTCTCTCTTGAGCCGCTGCCTTGTTACCAGCTGGAGCTTGACGCAGCTGTGGCAGAGG  
TAAACTTCGAGATGATCAATATACACTGGAACACATGCATGCTTTGGAATGTATAATTACCTGC  
ACTGTGATTCATGGTATCAAGACAGTGTCTACTATATTGATACCTTGGAAAGAAATTATGAATTTAA  
CAGTAATGTTCTGGACACTGCCTTANGAAAACCACGAGAGGTGTTTCGACTTCTCATTAGATTGACA  
GCATGTGACAACCCGCTTTGNCATCTATCCATTTCTCATCTTCTACCTGGGTTACCTTGTGAGAAGG  
GAACTGGAAGATTGTTGTCATTGNAACAAGGGGAACCGTGNAAAAACCGCTTNTGAAAAATG  
GGAGAATATGTTTAATGAAA

SEQ ID NO: 878 ACAACCACCACTCCTGTTCCCTTCCATTTTTTCTGGCCTAGTGTCAGTGCCAGGT  
CCTTCTGCCACTCCTACCGCAGCCACTCCTACCCAGGACCTACACCACGGTCCACTCTTGTTCC  
AGTGAAGCATTGTCTTCTACTTCTGCACCTTTCACTAGCCTCCCCTTTTCCACAGCTCTTCTGCTG  
CTTCTACCAGCAACCCAAATTCTGCTTCATTGTCTATCAGTTTTTGCAGGGCTCCCTTTGCCCTTACC  
AACCAACATCCCAAGGCCTATCCAACCCGACTCCTGTAATTGCTGGTGGCTCTACTCCCAGCGTTG  
CCGGTCCACTTGGTGTGAACAGTCCCTCTTTTGTCTGCGTTAAAAGGTTTTCTGACATCCAATGACAC  
CAATTTAATCAACTCCTCTGCTTTATCCTCTGCTGTCAAAAGTGGGGCTGGCTTCACTATCTTCTC  
TACTCTTCANAACTTGAAGCTTTNTGGTTTANCCCCCTNACAAGTGGCTNTGCCCCATCAGCCATTC  
CTACCCCANAGAGGACTTCAATCCAAGGTGGCCCTTNTTCCAAGCCTGTGCTTNTCCGGGGNTAA  
CTCAACTT

SEQ ID NO: 879 ACCTTTTGCTAGCATAGCCTGGGAAGAAGTCACTGAAGGAGATTTAACTGAA  
GGTAATACAGCTGAGGAATTTGCTCCAGAAACACTTTCGCTGGGTTTAGGCTTCTTTGGATTAAACA  
TTATTTACTTCAGAAGTAGAAGGACGTTTTGATAAAAATGAACTTTCATCTACCTGCCATAGAA  
AGTGTGGTCCCCCTTGGTAATGTGATAATCCTGTATCCTGGGTCAAGTCAAACAGTAACTCTGAT  
GACTTATTAGACACAACTCGTGATGAGTTTGTATATAACCAGGTTTAGAAAAAGGCTGAACAAT  
AACAGAACTATCTGAAGATCTAGACTCAGGATGAAAATTTATTGGTGAGGCAGGCACTGGATTG  
ATGTTGGATTCTGTAGTGTGACTGTGAGGCTTGAATGCAGCAGTAATACCTCGACTGAGTGCCC  
ATTCTTAAATCCCCTTGAATATGAGCTGGGGTTAACTCTGTTCAAATATTTGAAATGCTGGTTTTG  
ACNTGATATCTGCCCCACAAAGATTGGCTCATATCATCGTCATCCTCAAATTCCTTTACTTTTTNTG  
GCATGGTT

SEQ ID NO: 880 GGTACTAATCTCTCTGAATTTGTGTCATGCGGAAAATTGGAGACTTGGCTTGTGC  
TAACATTCAGCATCTGAGTAGTCGCTCCTTAGTGAATATTGTTAAAATGTTCCGTTTCACTCACGTG  
GATCACATCAATTTTCATGAAGCAGATTGGAGAGATAGCTCCTCAGCGAATTCCTTCCCTGGGAGTT  
CAAGGTGTCATGCACCTGACTCTTTACTGCTCGGCCCTTACGCTTCTGAATGAAGGAGTAATGAAT  
GCAGTGGCTGCGTCTTTGCCTCCTAGAGTGGCAGACTGTCGAAGTAAAGATGTTGCCAAGATTCTG  
TGGTCATTTGGAACCTCTGAATTATAAGCCACCCAATGCAGAAGAATTTTACTCCAGCCTGATAAGT  
GAGATTCACAGAAAGATGCCTGAATTCAACCACT

SEQ ID NO: 881 ACTGAGGAAGACACCATTCCTTGACGGTGTCTAAGAAGCCAGGTGGATGTGT  
GTGGTGGCTCCAGTGGGTGTTTCTACTCTGCCAGTGAGAGGCACTCCCCTAGAACTCTTCAGGCG  
TAATGGAAAATCAGCTCAAATGAGATCAGGCCCCCCCCAGGGTCCACCCACAGAGCACTACAGAGC  
CTCTGAAAGACCATAGCACCAAGCGAGCCCTTCAGATTCCCCCACTGTCCATCGGAAGATGTCTCC  
AGAGTGGCTAGAGGGCATCTAAGGGCTCCAGCATGGCATATCCATGCCACGGTGCTGTGTCAT  
GATCTGAGTGATGACTGCATGCTGCCTGGGATTGCAGCTGAGGTGGGAGTGAGGAATGTTTCCC  
AGGAAGACAGTTCCACCTCTAAGGTCCGAAAATGTTCCCTTTACCTGGAGTGGGAGTGAGGGGT

CATACACCAAAGGTATTTTCCCTCACCAGTCTAGGCATGACTGGCTTCTGAAAAATTCCAGCACAC  
 CTCCTCGAACCTCATTTGTCAGCAGAGAGGGCCCATCTGTTGTCTGTAACATGCCTTTCACATGTCC  
 ACCTTCTTGCCATGTTCCAGCTGCTCTTCCAACCTGGAAGGCCGTCTCCCCTTAACCCAAGTCCTN  
 CTCAGGCTTTGGAGAAACTTNCCTTAACGTACCTTCTTCATTGAGCCTTTTTGATCACTTCATCC  
 CTTTTCTAACCTTCCCTTCCCCAACCTTAATGNATNAAATTGCTTTTTGAAGCTTAACANTCACA  
 ATTTTGAATTGGAACCGTAATT

SEQ ID NO: 882 CCGGGCAGGACTTTTTTTTTTTTTTTTTTTTTTTTNGGNTAAAAGGANAATTG  
 TCATTTATTTTTGGTTTTGTTTGTACATTTTAATAGGATTGCTCTANCAGCTNNGTTGANAATAAN  
 ACAACACAGGGAGACAAGCNTAAAAGCCGGAAGGGGAGGTAGCTGCAAAAAGGCAGGAAAAAGG  
 TGAGGCGGGCTGGTCCAGGAGGNGGCCANNATTNNGGGGTGAGGCAATGGGAGTAGCAAAAGCA  
 GATTCTGGATGCTNTNACAGTAAANCCTGTAGGACTTGCTGATGGATCAAATGTGGGTGTGANA  
 AGTCAGGCAGGACTCCCAAGGTTTGGCCCTGNGGTGATGTTAATATTAAGTGTCCNCTTGATTGA  
 AGGATTNACAGAATTNTTCTAGGNGNGTCNNGAGGGTGTCCCAAAAGAAATTAACNTTTTGAT  
 TCAAGNNGGACCCGGAAGGCTTACCCNCCCTCNATTTGGGTGGGCCCCATNTTATCGGCTTCC  
 AAAGNNGCTTAAAAAAAACANANTTNTGAGNNTTTTGGNNTTCATTTTTTGGCGGAACCTGGGGC  
 NTNACCCCC

SEQ ID NO: 883 ACAGACAAGGTCTATAGAATGTGGTAAAAAAGTACTGACTGCAACACAAGGCTTA  
 TAAAAATAGTAAAGATAGTAAAAATAGCTTATGAAGAACTACAGAGATTTAAAATTGTGCATGACTC  
 ATTTGAGCAGCAAAAATAAGAACTCCTAACTGAACAGAAAATTTTTCTACCTAGCAATGTTATTCTTG  
 TAAAAATAGTTACCTATTAATAACTGTGAAGAGTAAAGAAAAGTAAAGCCAATTTATTATAGTCACA  
 CAAGTGATTATACTAAAAATTATTATAAAGGTTATAATTTTATAATGTATTTACCTGCTCTGATATA  
 TAGCTATAACCCAATATATGAAAATCTCAAAAATTAAGACATCATCATACAGAAGGCAGGATTCC  
 TTAAACTGAGATCCCTGATCCATCTTAATATTTCAATTTGCACACATAAAACAATGCCCTTTTGTG  
 TACC

SEQ ID NO: 884 GGTACCTCCAAACAGAGATGGAAGCTACACTGCAGTTCCTCAATACTACTTCA  
 GCATAGAGCAAAAATGTGAAGCCAATTAACAGAGAAATCATTTTTTGGCATTATTAGGCAATCAAA  
 GGGGTAACTAAAGTGAAGTGTGGTTCAGAAATTGAGAAATCTTTTTCTTTTTGAATAAAAAAAG  
 GAGATGAAAAAAGTCCACTTCTTCTCAGTGGTACTGTAGAAGATGTCTCTTACTAAAAAGGGGT  
 TTTCTACATTTTAAATGAGATTGAGGCTATCTTAGGGAATGAGCATTGTCTTTTCATATGATTAGT  
 GTCTACCCCAAGAATAGTTCCATTGATGAAGATTTTCTATATTTTTTCATATCTAGCTATGCTATTT  
 CCTCATGAAAGTCCAAGACTTTTTATGACTGTGGTAATTTTAGAATATACATGAATGATCTTTCAG  
 AGTCACAATTTTGCCATATCGTTAAAAAACTTATCCCGGTTCTAGCCTCTGNATTAGCCCTCTC  
 CTGNGCTATCCTAATCCNTAGATTAGAAAGAAGAAATCTGCTTNTGGNGCCCTNCAGCCACCAAC  
 AGA

SEQ ID NO: 885 ACCCTTTTTCTTTTTCTTTTTTTTTTTTTTTAAGTATTGTTAACAATCCTTTGG  
 AAGTCACTACTGGTCTTTGTGTGCTGCTTTTTAATAATTGAGTTATTTTGAGCTTGCCAAGTAGGAT  
 CTATTGCCTGGACTAAAATTTATTTCCTAATCTTCTGATGACCAAGAAAGGAAAAATTAAGTTTGC  
 AGATGTGAGATGAAATATAGCCAGTGAATATGCATACTGATTCTGAATGAAAGGAATTAACCTTT  
 CAGTCAAGAAACAGTCTGCATGCAGTAAATTGAATTTTCTGCAACTGGAATGATTTGTTAATT  
 CTTCTTTGAACACTGCCCTTTCTCCAGTAAAGACACTAATGATTTGCTAATATTTTTTAAAGAAATC  
 TGNTTTTTAATTAGTTAAGCTCAGACTTCCTCTTATTTTTTATCCTAAAGAAAAGTCTAAAAGGG  
 AATGATATATCAGTACC

SEQ ID NO: 886 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGTTATAAAAGCCCTTTTATAAAGCC  
 ATTTTTAAACAAAACAAAANAAAAGTTTACAAAAGAAAAAGAGATACAGAAAAAGAAATNACTTG  
 CTTTCATNTGTCCAAAAAGAGAAAAAATNAAGGGGACATGCCNACATGCTCAACAATAAAGG  
 CTTCTTTNCTTATTTTTTTAATACAAAANACAAGCNATGGATNCACATACTTAAACAGAGCTTA  
 GGAGCNGACNCNAGTCCCTNGAAACCNNTTNAATAAAAGCAAAGCAGGAGTTTGTTTTTCTTTGN  
 CTATNCANATGCNTACAAGAGACTGGGATTTGTAAAAAATTNAGTGGTCNCAAAGACCATNACAC  
 NATTCTACCAATGCATGTTGCATNTGTAATTNCCGAACATGGTCAACAAAANATCTTTGTTCACTTTT  
 GCCCCCTTTNATTTTNNTTTAAATGAANAAAACCTTTAAANAAAAGNNGGTTNCATTTTAAANTTT  
 AACTTCCTTAATANCTANGNNGNACNGCCTAANGGCGAATTTTTATCACACTTGCGGGCCGTT  
 TCTANTGGNTNCCACNTCNGNC

SEQ ID NO: 887 ACCAAGGAGCTCTTCTTTATTTATTTCCATATGGCCCTCAGCAGCTTTTCATCTG  
 GAGAAGCGAANCCAGGATATTNCTTTCTGATGAGCTATTTTCCCATATATACTTTCATGTNTGC  
 TTCANNCTACTNTNNNCCCATNCTTCANNNTTCCAAGACTNANGGCATANCACTGGNAGAATAN  
 ACGCTCNTNATCGGNATTCTGCCTCCTTAAAGGGCACCTATCTGTTTTGATGGTGAGCACATNTGT

GGTCTTGCANTATCGCTCTCCTACAAGGTTAATTAATTTNTTTTTTNCGTGATCATCTAAANTTANA  
CNGGAAAGCTCNCTCTTAAGACTACTACTCGTGGCTCTGGGGTCCGAACAGATGGTCCTGATGAA  
ACATAATCAGTGCTGNNNATTTCAATTGGAAAAATGCTNCTTACATTTCTNGNACTGTCAGNGCGGN  
TGGCCCTCAATGCAAAAATNTTANAGGCTNCACATTGGCTNTATAATTGNCACCGNNGTCAAAT  
GCCTAANANTGNNAATNTTTAAAAATTCGGGAATTTCTCTTGGC

SEQ ID NO: 888 ACTTTATTTTTNTTTTTTTTTTTTTTTTTTNGNGNGTTAATTCATTTTAATAGTAT  
AAACCTCATTTAGGTAGTAGTATTAAGCCACAACANTAATGCCNCATTGAAACAGCATTTAATAA  
AATGCATAAAGCTNATTATGCACTGGAATACTCTNTNTACAAACACAACAATGCAAATTTCTNTT  
AAGACTGAAGACATTGATTAGATATAAAATTCAGTTTAAAAAGAACATGCTTTTTTTTAAATGCCN  
TNNCATAAACAAAGNGATTTNACAGGGANAAAAAGCTGTTTAAAGCNGNNNTTTNAACAGGT  
TTTAAACCTGGCATTAAAATGTAATGGCAAAACCAATATTTTTATNTTGTNAANGGCAAAGGTTT  
CANAAACGNGCCCAAGAAAANNTAACNCCCAAANTTTTCTTTGATNGATAGATGCCTTTAAAAGG  
GCTGGTAATGCAGTTACATTTTAAACANAGAAGTTCAACTTCAGGTAAAAACTNGGGTTGTCCCTTN  
GGCGTGACCACNTTNGGGNGAATTTCAACNCCTTNGGNGCGTTTCTTTNGNG

SEQ ID NO: 889 ACTTTTTTTTTTTTTTTTTTTTTTTTTTGGNGACAGTTGTATTTATTTTTTAAGTTA  
CAATAAAATGCTCTCAAGTCCTTTGAATGTTCCAACAAATTCAAAACCTTCATTTTCTGAATGTTT  
CATAAATGCGAACTACCTGTTTCGCAATTGGTAACCTGCTGCTGTATTTTCATGTCTTAACGGCTATTT  
GAGGTTTCATTAACAACATAGAAAGCCTTGAACCTGTATAACCAGCTAGATTCCTTAATAATTAGTCA  
CTAGAGACAGGCCCAAGACAAATATTGGGCAGGAAATCAGTTCTCACTGAGCCCGGTTTCCATGT  
AAAATCTCTGTTGTGGTGGGCATAGGTGGCACCCTCTAAAGAAAAGAGGTCTGTTTTTTGTTTAA  
AAAAGTTTGTGGGGAGGAAAGACATCTGTGTATCACTTCAAATATTGATTTACTGCTAAACATCA  
CTCTGAATTTATGATGTGGATCTAACTTCATACATTTATCGGCATTGTCCAAAATATTTATTCTTA  
ATGGGAAAAGNCNTTAATNTTCNAATGAAGGGNCNCATTA

SEQ ID NO: 890 GAAGAAGCCAAAAAGAAACGAAATAGATGCGGAGCCGCCAGCTAAGCGGCA  
CGCCACAGCAGAGGAGGTGGAGGAAGAAGAGAGGGACCGATCCCAGGCCCCGTTTGCAAGGGA  
AAGTGGAAAAATAAGGAACGGATTCTCATCTTTTCTTCCAGAGGAATAAATTTTGAACAAGACA  
TTTAATGCAGGACTTGAGAATGTTGATGCCTCATTTCTAAAGCAGATACTAAAATGGATCGTAAGG  
ATAAGCTATTTGTGATTAACGAGGTTTGTGAAATGAAGAAGCTGTAATAAATGCATCTATTTTGAAG  
CTAAGAAAAAACAGGATCTCTATATGTGGCTTTCAAATTCACCTCCGGGACCATCTTGCTAAATT  
NCCCTGTCAAAAATATTCATCCCCCTCGCTTGACCTGAAGAATGACTTGGAACCTGGTTGGAAAGGT  
TTTCGGCCCTTTTGGCTTTTGGACCTTNCCTTGGNTGAATAACNCANTTTTGGCTTGGTTAAAAGAA  
CCTTTTATTTTCAATNTTTATNCCTCGGCCGGACNCCCTTAAGGCGAATTCNACCCACTTGGNGGC  
GGCTTATGGTTCCCACTTCGNCCCACCTTGGGGAACANGGCCAACNGTGTTCTTGGGNAATGGT  
TTC

SEQ ID NO: 891 GGTACCAGCACCAGCCCCTCTGAAAGGAAAAAGTGATGCTCATGACTGTCCAT  
CTCTTTTCAAAGCTTCCAGTCTTTGAAGCAGTGCGTTTCCAAATGCTTCTCGGTAGGCGGGGCTGA  
GTGTATCCAACAAGGAGTAGACTGTTTCATGGTAGGGAGTCTGTAAATGATCATCTACCTGGTCAA  
AAGCATAGCCTACCACCTTGAGCCCTGCTTCAGTGAGTTCTAGGCAATATCTGTTTCTTTCTTTGT  
TTCCACATTGATATAGGCCACATCATCCGCACACCGCAGGCTTTTCGAGACAAACATGTTGTTAAC  
AAGCAAAGAGAACATCATTACAACCTGCTTCAGCTTCGAGCCTCATGTCTTTTCATGTCAAGTTCCCT  
CAAAACCGTTTCAGCTCTGAACCTTCTTCAAATCCTGACATACTGNTTAAGCTNCATGGGAATACAG  
CTGTTTCATTCTTGATGATGCAGAAAGACAAGAATCCAGTACCTGGCCGGCGGGCGNTCNAAAGG  
CGAATTCACACACTGGCGGGCGTCTANTGGATCCGGCTCGGACCAACTTGGNGTAAACATGGCAA  
AGTTGTTCCGNGGNAATGGTNTCCGTACAATTCCCCCAAANACAGN

SEQ ID NO: 892 GGTACTACGTCAGCAATTTCTCCAAACAGCTGCTCGACAGCATATGGCACCA  
GCCATTTTCAATTTGCTGAGCATCGGCCAAAGCCTGTATGCGAAAGCCAAGGAGCTGGACAGAG  
TGAAGGAAATTCAGGAGCAGCTTTCATATCAAGAAGCTGTTGAAGACCTGTAGGTTTGCTAAC  
AGTGCAATTAAGGAGTTTCGAGCAGGTGCGGGGACACTTGACTGATGAGCTCCACCTGTTCTCCCTT  
GAGGACCTGGTCAGGATCAAGAAAGGGCTGCTGGCACCTTACTCAAGGACATTCTGAAAGCTTC  
CTTGACATGTGGCTGGCTGTGAGCTGTGTCAAGGAAAGGGCTTTATTTGTGAATTTGCCAGAAT  
ACGACTGTCTCTCCATTTTCAGACAGCAACATGTAAGAAAGATGTTCAACCGTGCCAGGGCTTT  
GCTTTTACAAACAAGTGCTTTTCAGTCCCTCCGAGTGCCCCCGGTGTGCGAAGGATCACAAAGCGAN  
GAGAAAACNTCTTGAAGGGGGGGCCTCTGCAGCAACATGATGCCCTGAGTACCTTGCCCGGCCG  
GCCGTTCTGAAA

SEQ ID NO: 893 GGTACGCGGGGCATGCGCGTTTCTCTGCATGGTGTGCGTTCTCGTTCTAGCT  
GCGGCCGAGGAGCTGTGGCGGTTTCTCTAATCCTGCGAATATGGGTAGTGCTTCGTTCCATGGAC

GTTACGCCCCGGGAGTCTCTCAGTATCTTGGTGTGGCTGGGTCCGGTGGGCATACCACTGAGATC  
 CTGAGGCTGCTTGGGAGCTTGTCCAATGCCTACTCACCTAGACATTATGTCAATTGCTGACACTGAT  
 GAAATGAGTGCCAATAAAATAAATTCTTTGAACTAGATCGAGCTGATAGAGACCCTAGTAACAT  
 GTATACCAAAATACTACATTACCGAATTCCAAGAAGCCGGGAGGTTTCAGCAGTCTGGCCCTCCA  
 CCGTTTTACCACCTTGCACCTTCATGTGGCTCTCCTTTCCCTAATTCACAGGGTGAACCAGAATTGG  
 TGGTGTGTAAACNGACCAAGAACATGTGTCTATCTGNGNATCTGCCCTTCTCCTGGGATACTANG  
 AATAAAGAAAGTGATCATGTGCTACGTTGAAACATCTGCCGGTANAAACGTATCCATGTCCGAAA  
 GATCTGGTTCATCTNTAAATACTTCATTGGTCAAGGGCGGNTNTNAA

SEQ ID NO: 894 ACACAATCAGATAACAAGGTTTCAAGCTCTAGTCAAAGACAAGATGATATTA  
 TACCTAAATAACAATTTCTAAGGAACTTGAATAACAATCTCTAAGGACTCTTGTCACTATATGAAG  
 AACTTAAGTTTATGACTCCACACTCTTAATTAAGGACAGCCACTTAATCAATTTCTGCAGTGA  
 ATTATTAGACTAACTTCCTAGCATCTGGAGATTACAGTATGTTCCACCTTGAAGCTATCACAATCA  
 AGACCATTACCTGACATGCATTCAACAAATATTCGATGACAAACAAAGTATAAGGTTAATTGTTAC  
 AGAAAGTTAGGTCTTCTAACACCTAAGTGTTAAGTGGACTCAATGTTGAACCTTTTTGAGTGCCAA  
 CATGACACTCAAAGGAAATGCTCATGGAACATTTTCAGATTTTCAGATTTGGGATGCCCAACTGGTA  
 AGTATAATGCAAATATTCCAACTTTGAAAAAATTATGAAATCTGAAATACTTCTTGTCCCAAGC  
 ATTTTTGAAAANGGATACTCAACCTGTCTCGNCGCGACACCTAA

SEQ ID NO: 895 GGTACCTTGCAGCCAAGGGAAAACTGAAGAGCCAAAACACCAAGCCTTATCT  
 AAAATCCAAGAATAATTGCCAGAATCAACCACCTTCTAAATCTACTATTAGACCCAAAAATGATG  
 TTACCAACCATGTTGTTTGCCTGTCAAACCTAAAAGGTCCATCAGCATTAAGTCCAGCCAGAC  
 CACCTAATACTGCAGGGTCCCAGAAGCCGAAGTTGGAGCCACCAAACTTCTGGGCAAAAGGCTG  
 ACTTCAGAATGTGTTTCTTCTAACCCATACTCTAAGCCTTCTAGCAAGAGTTTTCAACAGTGTGAA  
 GCTGGATCGTCCACAACAGGAGAAGTGTCAAGAAAACCTGTGGGGNCACTTAATATAGAGCAATT  
 GAAAACCTCAAAGCAGCAGTTAACAGATCAAGGAAATGGNAAATGTATAGACTTTATGAATAATAT  
 CATGTTGAAAACGAATCTTTGGATACTTTCTTAAAGAACCAACCAAGAGAAGTGTGNTTCGTTATNT  
 TAACAGAACCTGAGAGGAAGNCCGATCTTAANTNTTACCCGAAGTAAGCCAAAACCTNGNTTTTT  
 TTATTAACCCANAACAGTTTAGTTCTAACAAACCTTGGCNAAAGTTANTTAATAGGGNGGTTTGN  
 AA

SEQ ID NO: 896 CGAGGTACTNCTTCCAAATGACGAATTTTCTGCTCCAAATAATGGGACAAAG  
 GGCATCATCATGAGACACAGTCAATTTACAGTGATGGACTCACACAGGTGGATGATCTGAGGC  
 ACAAGCTGTGATTGGTCTTAAGAAAATGAAAGGTAGAGTTGGGTTCAAATCCACATCTCACGAT  
 TTCTGGTCTTATTCATCCAAAAAGTTGAACTGCTTCATATCAGNGTATTAGACACATTTTAAATCT  
 TTCTCAAGAGTTGCTCTGGGTGCAGCATGATAAGGAAAACAGGTATTTGGCTTGATAGCAGATAT  
 TCAAGGTATAGGTGCCNAGCTTATGGTAAAGCTTTTACCCGTAGGATATCTGAGATCTGNTGCCAC  
 AAAGAGGACTATGGNCCTATGATGACTACTATAGTTTCAGTTTCAACCTTTTTGGGGATTAATTTGA  
 TCATAACCTTTCTTCTCGTTNCATTNTAAGACAGGGAATTTAAANGATTCACTCAAGAAGGTTTT  
 CCCANTAACCAAGNGATGGTCATTCCATCTGNAATGGGTCCCTAAATAAACTGGCTTCCCCTTTCA  
 ACAACTCTTGNCNCCACANTCATTTGCTTTTCCAANTGGAACAAGTCTTTGGAATTTGAAAAAAC  
 TGTTTTGGNNG

SEQ ID NO: 897 GGTACTNGGGGGACTTAAGATGGCGGCGTTTGCACGGAGTGCAATCACTGCG  
 TCCTTACGGGGGTTGCAAGGCGTCCGAAGTATGAGTCCACTAACAAAAGTCCAGAACTCGCCAG  
 TTAATAGTATTGTGTCTCTTCAAATATCGGAGAATAATTTCTTTCTCGCTGATCGCCTAATCTTA  
 CTGACGAAGCTTGGAAAGTTGCAGAAGGTTGGAAGTGCAATGGCGCGATCTCGGCTGACTGCAACC  
 TCCACCTCCTGGATTCGGGCTATTCTCCTGCCTCAGCCTGCCGAGTAGCTGGGATTACAGGCATGT  
 GTCACCGCACCCAGCTAATTTGTATTTTTTAATAAAAGATGGGGGTTTCTTCATGGTTGGTCAAN  
 GCTGGGTCTTTGAACTCCTGACCTTAGGGTGGATNCGGCCACCCCTNAACTTCCAAAATGTGNTT  
 GGGATTTCAAGGCGNTAAGCNACCCGNCNCCGGGCCAGAACAAATTANTTTAAACTTTCTACCC  
 TTTGGGAGAAAGGCAAAAANACTTTGATGAAAATCTTNNTAANGACTTNAACATTTTNCNCCNNA  
 ANTATATNTTAAATTANTTAATTTCTTNCNNGGCGGCGNCTNAAANGGGGAAATTCCAACAN  
 NTGGGGGCTNTTTATNGGNNCCAACC

SEQ ID NO: 898 GGTACGCGNGGGAAGATGGGAAGTAAGAGTCACATATCAAAACTACCTCCAC  
 TTTATTCCCTGAGCGAGGGTTTATGAAGTATAAAGGGGTGGGAGCCCCGAGGTGAGCGGGAACGG  
 TGCTGCTTTATTTGAAATGTTTCTTACCTCATCTGTGCCCCAGTAGGGGGTCCAGCCTCATCTGT  
 CTGGCTTGGCCCTGTGTTCTCTGTCCCCGTCTCACTCTGCTGCTTCTGGTCCCCAGGTGCTGCTTG  
 CCACTCCAGCTGTACATTGAACAGTTTCAATTCAACTCTTAATGCTCCTGCTTCCGAAGCCTGCC  
 AATTTCTTTTTCTTGGCCTCTGGTTTTTTTTTTCTTTCTTTTTCCCTTGGTTTTGTAAAAAACTCANA  
 AGAAAATCTTTCTTATGGCTTCTTTGGTGAAAATGGAAATGGAAAAAACTTAATTTTTTGGATT

TNAAAAGGNAGGGCATGCCTATAACATTTCTTCTATATAAGAACTGCNTTGNTNAGGGGGCCCTCT  
GGTNGGGCCTACCTCATNAAGGGGTGGATNTTCCCCCTCCNTTTAATCTNGCGGGTTGGACCTAA  
TTGAAAAACCNATAAGGTTGTTGCCCCCNGGGGANGGAGTGGGCCTTGCCNTTNTTC

SEQ ID NO: 899 ACATGGCTCCATGGAGGTTCTCCAGTCGGTGTGCTGCTGCTGTTTTTCGAGCC  
TTATCTCGTCTGTGCTCCTCATAAGTGTGCTCATCAGTGGGCAGCTCATAGCGGCACAAGGGACAG  
GAATTTGTCTTGTAGCCAGGGCAGAATGCAGCTGGAATGGAAAAGGTGATGGCAAGGCATCTC  
AATGGCAGTCTCCTCCTCCAGCCTGAGAGCCTCTGATGACTGTCCTGGGGAGGTTCTCAACCACAG  
TCTTGGCAGCTGGTGGAGGCAGGTGGTGGTCCCAATCTACTACCAACCCCAAGTCTTCAAAGTCCA  
TCCTATTGAAAAGTGACCTTGGGAGCTCCAGCAGCATGTTGGTTCGCGTCTCCTGCTCAGGGTCCG  
ACGGCTCGCAGTCGTGTTTCATCGAAATAGGACGCCATGGCTGCCAGCCTTCTGACCCCGCGTACC

SEQ ID NO: 900 GGTACGCGGGGGGGTTCGGTTTCCGCGGTGGCCATGACTGCGGCCGTGTTCTT  
CGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCCCTTTATGTCTTACCATCGCCACCGAGCCG  
TTGCGTATCATCTTCTCATCGCCGGAGCTTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGT  
TTGGTTCATGGCAAGAGTCATTATTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTT  
TGGAGCGTTTGTCTCTGTCTATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAA  
AGCCAGTGAAGGTTTGAAGAGTATAAACCCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCT  
ATGTTTCTGGCTTGGGCTTTGGAATCATGAGTGGAGTATTTTCTTTGTGAATACCCCTATCTGACTC  
CTTGGGGCCAGGCACAATGGGCATTTCATGGAGATCTNCTCAATCTTCTTATTANCTTTTCATGAC  
GCTGGCATTATCTTGTGCTGATGTATCTNNGGCGATGGATTTTGTATGNTGTGANAAAAAANTGGG  
GCATCTTCTTTNGTTTNTNACCACCTNTGGGGTAAGCCAACTTCATANNTTNTATNTGGAANAA  
ACCGCGTCGCAATTTTAAACNTGGGCTCANG

SEQ ID NO: 901 CGAGGTACAGCAGCTTGGGAGTTCATTGCTGGTCTGGGACTATTTGCTTGGGC  
ACGTGTATAATGGCTGTGAACAGCGTTGACGTTTGTATTGGCACGAGACCTGCGTCGTGATGTAAT  
GCCAATATTTTACCTGATCATTCAATAACGATCATGGTGCGGGTGAAGAGATCATGATGAGTAATT  
ATGCCAGTAAACCACTGGGTGGCAGAGTCTTGTCTATATACTCTCACTCTGTATCCATTTAAGGAA  
TAAGGACCTGGTAGGGCTGAAAGGCATATCTTCAGTTAAAAAATCCAGTTGCTTATCTACAAGG  
AATTCTACTGCAGTGAAGTGAAGTGGGTATATTTCTTTCAACCAAGAGGTTTGAAAGTTTTGTTCTTG  
AGAATTTTAAAGTATCTTTAAATCTCAAAAAATCCACAACAAGGCTTTAAAGGNCAGTAAAGTATTT  
GCTGAAATTTCTATTTTGGTCTGGAAGAATATCTANTGNAAAAACCTCTTTAGTCTTCAAAAATCA  
CANGCCAATNGTNTTCAAACCAACANTTNTGAACTCCTTGCCGCGNGNGTCANTGTTTTGGAAGAA  
AAGCCNGACCCCCAACGAAGCCGNCCTTGTACCTAACAAGNTCCACCAGCCAAAGACAANTTTTT  
GTTGGNTTGGTCANCCNGGCANGATAAGGATACACTCCTTTTTTCATAANCCTG

SEQ ID NO: 902 ACCACATGTCCACCAGGAAGGAGCTGAGGTCATGCTGAGTTTTAGGCAACGC  
ACTGAGGGAGTCCACAATATCAAGTTTTGTGTGCTTCAGCAGCTCCCCCGAAGGTAGGGGTCCTT  
GAGAATTTTATATTCCTGCTATCCTTGAAGAACTTGCAAGCTGCTCGGGGCTGAGTCCTCAGG  
ATTGAAGGGAATGCTCAGGGAGAAGCAGGCCTCATCGTGGTAAGCACTAAGGAGACCCTGTGCGAT  
CTCCAGAGTCATAGATCAAGTAATACTGCTGCAGGAATTGCAGGAACAGATTCTTCAACATCTCA  
GATCCAAAGAAGCTTCTTACAGGTTGGTAACCTCTTGTGGGCTTCANTACCCGGGGGGGACCCCG  
GCTGGAAGGCANGGCATCACTATGGACAACCTGGCAGGATCCACCTNAAACGTCATNGACTCG  
TNCTGGAAGTTCAGAACAGNCCTACGNGGNGNTTATTTGAGGCCNATGGGAGAACGAATTT  
GAAAAGNCCCTTTTATGACAAAGCGCCTTCAAAATTGATCCNAGGAGAATCTGACTTGNTTGGTT  
CCANCAATTTTTTGATAAGAGCGTNTCAAAAAAAGGCCAGACTTTAAATAGGNCTGNTTNTTAA  
AAAAAAAAAAAAAAAAACCTCGCC

SEQ ID NO: 903 ACTGCTGACATCCAAAACCTATGTCCCTTTTAGGGTCTACTCGGAGAAAAATTGC  
GGCATTCAAAAAGTCAGGTGACCAGGGTAGCCACATTTTTTACAGCCTGCTCTGACTGTCTTGT  
TGCAACCTGGGACTGCCATGACGGTGGTAAGAGGGGTAACTCGAGCCTCTGGCTTTTGAAAAGGC  
GCTTGCTTCCCGCCAGTTGTGAGAACAGGCACAGTCAAAGCGGCGTTTTCTTCCCCCAGCCCCG  
CGTACCTCGGCCGACACGC

SEQ ID NO: 904 ACTTTCTGTTCTTTGGCACATTTTGCCAGCGGATGCAACTTCTATCCTCAGTC  
CAGTTCATATATCTCAGGCAGTGATCCATTTTGTATCAGCCAGTTTCCCTTGTAGGGCCGCTAGAC  
CCGGCTGGCAGCCAAGAGCAGCACTCTCCCCACCTTCAGCANAACTTCTCAGCTCATGTGTGTGT  
TTTAAAGCAGTAGGCAGCTTCCATCANCAATGGAAGTGTCCCCCACTNCAG

SEQ ID NO: 905 ACAAAGTGACATACAATTGGAATCCATTTTTGTTGTAAAGACATTGTTTTTC  
AGACTTTTCAGATCAATTAGAAAAATGTCATTGCTTTAAATCATAGCTGTTCTGTTTAAAGAGGAA  
TTGAATTTAAAAATGAGAGAGTATTAATACTCATGTGGCAGTATCCTGGTCTTAATCAGGTATTGC



AGTGTTCAAACAAGGTATGGACACATCAACATATTCACCTTTTGGAAGGACTAGTGACTTGGTTAA  
 AAACATAATGAGTGTGTTGCGTGTGGTGTCTATGCGAGCACCCATTCTCAATTTTTTTTTTTTTTTT  
 TTTTTTGTGTTGAAGCCATTGATTGATATTCTTGTTCAGTGTCTGGTCCCTATAAAGTAATATATA  
 TTTCCTCTTTAAGCTCTGNGCCTNAATATGCAATCACTTTGATATAANNNTAAATATGAAGACNTN  
 CCAAACCTGGNAAATNGGATTNTTTTTTCAAACTTTNTGGGGAAACATATGGGGGATATGATANAC  
 TTATGGGTTTTNTGTTGCCGGGGATNGGGTTG

SEQ ID NO: 906 ACTGAAGATTATTGCTTCTAGGGCATTTTTAAACAGCACCATTGTATTGTTGA  
 ATGTTTATGTAAGTATGGCTTTTCTATAATGTAATTTTTGAATGTTTCAGGTGTTACATTTCCAAAG  
 TTTAACTTTTAAAAAACCATCTTCTGATCCCTTTTATTGTCTGGGCCATACAATCTATATTACATAG  
 GTGCCAACATTTAATTCTTTTAAATGGAACATTTGCAGTTTTCCATATTGGTACAAAAGAATTACTC  
 TGATCGATATTAAATCGTATTGAAAACAAAATGGACTAAAAAGCAAATACTACTCTATGTTGGGG  
 TGGAGTGGGAGGAAAAGANTGACTC

SEQ ID NO: 907 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAACCGGANTCCCACTCTTGTGCGCCCAA  
 CTANAGTGTAGTGGCACAATCTCGGCTCCCCACAACCTCTGACTCCAGGGTTCAGGTGATTATTTT  
 GCCTCANCTCCCAAGCAGCTGGGATTATANACACCCGCCAACACGCCAGGCTAATGTTTTTGTAT  
 TTTAATAAANATGGGGTTTTGCCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCAGGTGATC  
 CGCCGCTTNGGCTCCCAAAGTGCTGGGATTACAGACGTGAGCCACCACTCCCGGCCATAAAAG  
 GNTTTTNGCTGGATAATTTGTAACCTTTCTAATTGGGAAAAAATTCCTATTAATCACTTAAAAATTT  
 TTTTTNTGAATTTGGNGTTATTTGATTATATACAAAGGAGACTTTTTTTTTTGANATAACACTCA  
 AATNGTNTTTCTTTNTTTNGAAAAATTTATTTTTTATNTGAAAATAACAGTTG

SEQ ID NO: 908 ACAAAGGCTGCTTAAGGCAGTGCAGCCCCTTCTCAAAGTCAGCATGTCAATG  
 ANAGACTGCTTGATACTTGTCTTCNGAAAGCTATGTTTGCCAACCAGCTTGATGCCCGAAAATCT  
 GCAGTTGCTGGGTTTTTGTCTGCTCCTGAAGAACTTTAAAGTTTTAGGCAGCCTGTCATCCTCTCAGT  
 GCAGTCANTCTNTAACTGTCNGTNNGGTTCATGTGNATGTACACAGCNATTACANTTCTGNCGCC

SEQ ID NO: 909 ATCCACCTCCGAAACCCCTCGGCGGCGTCTTCTGTGTGGCCCGCTGCNNGA  
 TTTTAANCTTGACTTTGGCAATTNCNANGGCNAANCANNNCAAACTTGGCGTGNANGGATNNCNA  
 CNATTTNTNAAANTCCTGNNNATNAATTGTGGGGAGTANTATGGAAAATGAACANAAGCANTTTA  
 AATNCTCTGGATGAGCNANAAGGGGTAAAAAGNNGGAATGNCTGTTGTAATATAAGTTAAAGTTGC  
 AACTCATGANGGAAAAGAAATAACCTGTGCAANTNATCTGATGACAAATNACGAAAGTGCTCCCN  
 CATGCCCACAGTATAAAAAGATTATTNGCATGGGTGCAAAACAAAATGGTTTGCCGCTNGAGTAT  
 CNTGATAAATTTTAAAGCAATACATCCCTATGACTTACAGGAATAGGTCTCTAATAAATTTGANAN  
 TCATCANAAAGGGGANACNCACTTTTATAAANACNAGATATTTCTAAGGGGTNTTCTNTG  
 ACCTNATATTAAAAATTTNTAACACTTGA

SEQ ID NO: 910 GTACTGACTTAAATTTGGAATTTACTAATTACTGGGGATACTTTAGNGAGTCT  
 GCATATGTGTATTATTAATACATGTTAAACCATACTGCAATATAACAAAAAATATACTGACATTTT  
 TCTNCANAGAGTAATGACTGTATTCAAAGTCTGAGGGAATGACAAAACGGGATGCACATCTAAC  
 ACTGATACACGGTTCTTCANAAAAGACTAGTTTCAGCTGTTTCCAGGTTTACATAAGATGATGGAA  
 GCAGTCTCTAATATGTTAATCAAGAAAAATATATGCAATTGCCAGCTACTACATATACAGAAAAATA  
 AAGATGGTGAGTCAAACAAAACATACAACTTGTAACCTTCTCACCCCTCCAGATATTTTGAACATA  
 TTTAAAAAATTTGAAGATAGGATATCACTGTGATCTTAAAAAGAGAACTTTTAAAAATAACACATT  
 ATTGGAATAGGAAAACCTGAGCTGCTAAAACCAAAAGTANTNAAACAGTTTTNCTGATGTCCTTAC  
 ACTCTTACTAAAACCCNAATTTAAACTACCANCTAAATTTATCTAAGTGTNGACTTTTAAATGN  
 CATTCTTTCAAACCTCCTTTATAATATTTTAAATTTTCTAANCNTAATNCATTNCTTAAGAGTNAA  
 CCCNT

SEQ ID NO: 911 ACGCGGGGTGGGGGGGGTCTGCTTGGCTTCTCGACTCGGTCTGTTTCG  
 ACAGCGAACATGTGCGGCCTGTCAGAAATAGGAAGGTTGTTGATTACTCACAGTTTCAGGAATC  
 TGATGATGCANATGAAGATTATGGAAGAGATTGCGGGCCCTCCCACTAAGAAAATTCGATCATCTC  
 CCGGAGAAGCTAAAAATAAGAGGCGATCTGGAAAGAATTCACAGGAAGATAGTGAGGACTCATA  
 AGACAAANATGTGAAGACCAAGAAGGATGATTCTCACTCANCAGAGGATAGTGAAGATGAAAAA  
 AAAANATCATAAAAATGTGCGCCAACAACGGCAGGCGGCATCTAAAGCAGCTTCTAAACAGAGA  
 ANAGATGCTCATGGAANATGTGGGCAGTGAGGAANAACAAGANANGAGGATGAGGCACCATTC  
 CAGGANAAAGATTCCGGCANCGATGAACATTTCTATGGAANATGATGACCGATAGTGACTATGG  
 CAGTTTCAAAAAANAAAAACAAAAGATGGTTAAAAAGTCCAAACCTGAANNAAAAAATNNT  
 ANAAGTCCCTCGGCCNNANCACCCCTAAGGCC



SEQ ID NO: 912 ACATAGAGAAGAAAATTTGGTTTTAGCAAATGACAGAGCCTTCAAAAATATT  
TTTGAATAATGTGAATCAACCGAAAACCTGGGGGCAAGGCAGAGGACAGGTTTTCTCAGGTTAAG  
AGAAAAACGAAATTTTAAAAACTTTAAAAAATACTGATAAATTCGGATCAAATTTGGGGGAATAA  
AAAATATTANAGCAAAGGAGTTTGGCTGGTTGTGTGTCATTATTTAATGATCAAAGTATAGCATTGTAT  
GCCTTATTACAGACTTGTGACTATAGGCTTAATGTAAAAAGGAATCTTTGCCAGATGTNGCTCT  
TAAGGAAAAAAGGGTTTTTAATAANAAATGACCTTTTGATTAGTATGGTGCCNAGTCACAGGGCN  
TATTTTCCTGANTNTTGGGGTGATGNCAAAAGGTAT

SEQ ID NO: 913 ACGCGGGCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACC  
TGAGTGCTACTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGT  
GGAATCATGTGCGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTANCTTGTCTTTGGCC  
ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTATTGTTATTTTGGG  
ATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGATTATGACTATACCAACGACCTCAGCATAGAA  
TTGGACACAGAAAGGGAAAAATATGAAGTGCCTGCTGGGGGTTTGCTATCGTATCCACAANGCATC  
ACGGCAGTGCTGCTCGTCTTGGAATTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTT  
CCAAATNACAAATAAAGCATCAGCAGTGCTCCCTTCTTGCTGTTCCACNCACTGGNGGACATTTT  
GCCATCCTATTTTCTTCTGGGNCCTCTGGGTGGCTGGGCTGNTGANCCTGGGAACTGNAGGAGCTG  
CCCATGTTATGGAAGGCGGCCAAAGTNGAATATAAAGCCCNTTCGGGCTTTTCGGGAACCTCGG

SEQ ID NO: 914 ACCAGCACATGAAGCCCTTCTACAAAATTCCTGACGGACTGGGAATAAAAAAT  
TCCTAGTGACAGCCCACTCCTTCTCAGGCAGGTGTGATTGTTTGAAATCTCTCCCAATATTGAGAT  
GAAACCTGCTTCCCTGTAACTTCCCTGTAAATCTGTGGGTCCCTGTAGCCACAGAGAAGGCAGCA  
ATCAGTAGGGGAAGTGCTATAAAAAATACTATCCCGGCCAAGCGTGGTGGCTCATGTCTGTAAT  
CCCAGCACTTTGGGAGGCCAAGGCGGGCAGATCACTTTAGGTCAGGAGTTGGAAACCAGCCTGGC  
CAACATANTGAAACCCCGTCTCGCTGGGTGTGGGTGGCTCATGCCTGTAATCCCACTTTGGGA  
GGCTGAGGTGGGTGGATCACAATGTCAAANAGATGGANACAATCCTGCCAACATGGTGAAACCCC  
ATCTCTACTAAAAATACAAAAATTANTNGGGCGTGGTGGCGTGTGCCTTTNANCCCACTACTTGNG  
AAGGCTGANCCANGANAATTTGCTTGACCCTTGGGNGGCANNAAGGTTGCANTTGAACCNAAATA  
TCACANCACTGGACTCCAAGCCTGGGTGACANATTGAAACTCC

SEQ ID NO: 915 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATNGAGTATTGTTTTATTAACCAA  
ACACAAAACCAAATGAAAACCTGGCTTANAATATAAAATTCCTATTTTTCAAAGTGAAAGTTTGA  
ANATACTANCTAAAGTTGATAACTTAAATAGNNGTAAAAGTAAATAACTTAAATTATGGCACCA  
ATCACAAGAAAGAAACAGAAAACAGGGGGACTAGGGATTCCGTGGTANACTTTTACTTTAAATAN  
AGCTATGCAGCANATTCTCCATGACTTGGCTTACATGCAGTATGTCCTATGGAAGTAAAATTTTCA  
AATACCTCCATCACCTTCAGTAACTCATATTAATAAAGTAAAAGCCANGTNTATAAAACAANAAC  
CCAGTTTATAAACATATACATCAATTGGATGCCAGGTAAACCAAAGGAGTTTTTAAAAAATATCC  
CCACCATACTTTTAAAGACAACCTTTCCCTTCCCTACAGTAGGAAGTACCACCATTCAAATTAGTT  
TAGCTATAAAAAAGTAAAAGGGGCCNCAACCTTTTTTAAATTAATGCCACCA

SEQ ID NO: 916 ACGCGGGGGCCATGGCAGCATCTTCCCTGACGGTCACCTTAGGGCGGCTGGC  
GTCCGCGTGCAGCCACAGCATCCTGAGACCTTCGGGGCCCGGAGCAGCCTCCCTTTGGTCTGCTTC  
TCGAAGGTTCAATTCACAGAGCACTTCATATCTACCAGGATATGTTCTAAAACATCCCTGAGTTC  
ACCACCTTGGCCAGAAGTTGTTCTGCCAGACCCAGTTGAGGAGACCAGACCATTCAGAGGTCG  
TGAAAGAAGGTGAATGAGATGATCGTACGGGGCAGTATGGCANGCTCTTTGCGTGGTGCACTTT  
TCCACCGCCANTGGAAGGNGACCTCTGAANACTNATCTTAATTGGAAATGAACTANACTTGCCTG  
NNGAGAGAGAATCCACTGGAAAAGNCCTGCTGGTTGGGGCAGACACTTCACGCTGTTGGCAAGCC  
ACTCTCGAAAGGATCTTGTGANTAAATNCACANCATTGAAAACAANTCTTGGCCAAAATCTTNTG  
AATCAGGAAAGGAAAACCTCAAAAANAAAAAT

SEQ ID NO: 917 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAACACAAGGGTCAGTTTCTTCAATTC  
ATGAGCAGTCANAACAGGANATGCTTAGGAAGAAATCGTGGCTGGTGCCTNTTCTCCATGCTCAT  
CCCATACCCCACTGACAGGATACCGCTTCCCTGAAGTTAAAAACATGCACCACTTCCGGTAA  
GGCTGGAGCCACAGAGGGCACCTAGCACTTGTCTTTCAAGGTCCCATCAACAAATGAAAAGGGC  
AGCCGAGGACAGATTGACAGAAACCCCACTATGTTGCTTTTCTTTTTTTTTTTTCTGAGGCAGAN  
TCTTGCTTTNTTGGCCANGCTGGANTGCAATGGCACAATCTNAACTCACTGAACCTCCACCCNTG  
AGTTCANCAATTCTCCTGCTTGCCTCCCGANACTGGATACAGCCCTGCACATGCCTGGTAATTTG  
TTTTTANTANAACGGATTTCNCATTTGCCAGCTGTTTACTGTGACCTAGGTACCCCTGCTACCTCC  
AAATGCTGGNTA

SEQ ID NO: 918 ACAACCTTCAAACATTCCAGTTTTTATAAAAAAGGGGCACACAATCGTGGTTT  
TGATCCCTTTTGTTTTTGGACAAATGTTTCTACAAATACAGATTTCAGCAAACCCAAAGGCTGCAA

ATTTACTTGCACAGTAATCTGCCAGCCCATTACTCCACTTAATCCAGCTGAACTTGAAATGCAAA  
 CCAAAATGTCCATGGTCATTAGCAATCATAGCAGGTAGAAAGGCTTTATAAGTCCATAAATGTGCTT  
 TGAAATTCACATCAAATGACTTTTCATAAGCTCATCTGGACAGTCAAGGAACTTTTTGCCTGTTA  
 CGATTCCGCAATTGTTGATTAGGATGGAAACATCGCCGACTTCTTTTAACTGGTCGGCTACTCTAT  
 ACACTCCTTCCCTTTGGCTGCAATCGCANGTATAGCGTGCCTCTTGTGGCTCCAGCTTCCANCCAT  
 CTTACATGTTCTATTCCCTCCTATTGATTCCAGAAACAAAACAGATCCAGCGGCAAACTGAAGGGT  
 ANACCTTCCAGTCCTTC

SEQ ID NO: 919 ACAAGCCACAAAAATGCTCACACAAGCGGTGGTTGGAACAAGCATACAACA  
 CAAATATTCAAGCTTTTGTCTGATGCACAGCCTATGAGCAAATTCAGCCCATGGTTTGAATCAACA  
 TGCCAGCTACTATTTGGCTGCTCCACAGTTAGCTCCAGACTAGAATTATTTAAAAAACAAGGCA  
 AAAACAAGAATGTTGGTCCAGTAGCTCCTAAACCTGTGAAGACTATGACTACTGCTGTGTCTAGTG  
 TGTAGTTTCTATTCCAGCTTGTGTTAGGCTTCCATGAACTAATTTCTTACTTTGCTTAATTAAGC  
 ATAATCAGTTCATGGAAGACCATTCTTCTTCTCTAAGTATGGTTGAGGACAAAGAAAAGGAG  
 AAAAATCTTTGCTGAAATCAAAATGGAACCTTGCTCATNAAAAAGAAAGTGTGTGTCAGGGGAAC  
 GCANATATCTATCCTATTCTTATCTCTACCAGTGAGGAAATGCAAAGGGCAGTGGGTCATGANAA

SEQ ID NO: 920 ACAAAGAGCCCCTCTACGAGAAGGACAGCTCTGTTGCAGCCAGATTTACGC  
 GCATGAGGGAAGAATTTGATAAAATTGGAATGAGGAGGACTGTAGAAGGGGTTCTGATTGTACAG  
 TTCATATCCCAGTTCTAGAATCAGTTCATTTCTAAGGAGTCTGTTCTTTTATTGGAACCAAA  
 ATCTGGGCACCAAGGTGTGCTCCCATTTCTAGTTGTTTCTGACCACATAACTGCTAACAAAGATGCT  
 TCACTCTGGCTACACTGATGTGAACTTTGAACTTTAGCAGAAGAGCTCAGCTCTAGAGAACAATGA  
 GCTCCTACATTACCTTTTTCTCAAAGAATAAGTAAGTCTAAGCAGAAAAAATATGCAAAGAAT  
 TTTAGTATGATGAAATAAGACAANCATCAGGCTTNTGACTGTAACCAACACAATATAGTTATAC  
 AGATCTGTAGAAGATCCTAAAATAAGAAAGTCATTTGCNGGGGTATCAGGGANATCTGNTGTATT  
 GCTTTGCCGCTACANAACACATG

SEQ ID NO: 921 ACGCGGGGCCCCGAGCTTGGAACCTCGTTATCCGCGATGCGTTTCTGGCAGCT  
 ACATTCTGCTCCTGGCGCTCAGCACCGCTGCCAGGCCGAACCGGTGCAAGTCAAGGACTGCGG  
 TTCTGTGGATGGAGTTATAAAGGAAGTGAATGTGAGCCCATGCCCCACCCAACCTGCCAGCTGA  
 GCAAAGGACAGTCTTACAGCGTCAATGTCACCTTACCAGCAATATTCAGTCTAAAAGCAGCAAG  
 GCCGTGGTGCATGGCATCCTGATGGGCGTCCAGTTCCCTTTCCATTNCTGACCTGATGGTTGNAA  
 GAGTGGAATTAAGTGCCTATCCAAAAAGACAANACTATNACTACCTGAATAACTACCAATGAAAA  
 NCGAATTTCTTTANAACTGGNNGTGGANTNGCACTNAGNTNNCAAAANCAAANTTTTTNTGG  
 TNGAAATTCNATTNCNGGGGAANTNAGGTTNATTNCGGNAAGNNCTNAAAACGGTTCNATTCA  
 AGAGG

SEQ ID NO: 922 ACATAAGCCTTGATATTCCATTTTGTGGCTGGTCCAAGGGGCAGCCTAACTCA  
 TCTTACAAGACGGACTCTAGGGACAGGTAGGTTGGATCCTCATTCTGACAGTGCATTTGTCTGAC  
 ACACGGGACTACCTCCCTAATGTATGCTTTTCTTTGAGAAAAAAGTAACAGATTTTCTTTACA  
 TCTTGGCACCTTGTAAGTTTTTTTTTTTATACAAAAGTCAATAGTTTTGACACTCCCCATTGTTA  
 ATCACTACTTCACTGATAAACTTGAAAAAGTGTGACCTGGAATTCATCATGCAAAATATTTACT  
 GCAGCAGGAGAAAAACATTTTTTAAACAACATTTTTTTCTTTCAAATGTATGAACTTGTTTAAAG  
 TAGCCAGGAAGGCAGTGGTAGGATAAACACAAGGGATAGGAATGTATCAAAAAACAGATTAACA  
 CACACGCACGCGCACACACACACACACACACACAAAAACCTGT

SEQ ID NO: 923 ACCTTTCATTGCAGGATTTCTGCTTAATATAACAAGCAAAAAACAACAACCTG  
 AAAAAATATAAACCAAGCAAAACCAACCCCCGCTCAACTACAAATGTCAATATTGAATGAAGC  
 ATTAAAAGACAAACATAAAGTAACCTCAGCTTTTATCTAGCAATGCAGAATGAATACTAAATTA  
 GTGGCAAAAAACAACAACAACAACAACAACAACAACAACAACAATCCCA  
 ATCTTCATGGGTAACTTTCTGCTCAGGGATGTAAGCTGACTCTAGACAAGATATTGT

SEQ ID NO: 924 ACTGCCGAATGTGTTTTTCATGACATACTTCGTAAGTCCAATAAGACTCAATT  
 CTGTAGGAACAACCTGCTTTGTTTAAATAGTGGCTCCAAAAGCTCTCTTGGATTAGGGCCTTTATAA  
 TCCTTTTCTTCTTCTCATCCCCACTTGTCAAGGGGAAGTATGCATTTATATTTTCTTNATGTGC  
 AGTTGTCTATGANGACATAATNATCTTCTTTATATAANACTNCNGNTGTGGGCANGAGAGAACTCG  
 GTGCCGGGCCAGNTNACTCGNAAAGGGATGTCCTCCCTGAGTTGATGAAGGGCTCGGCCGCTGC  
 GGACGCCTCCNGGAGGCCNATAANACCTGTAACCTCCCGGCCCGGCCCGGACCCANACTCCGT

SEQ ID NO: 925 ACGCGGGGGCAGTCACGGGGGAGCGAGGCCTGCTGGGCTTGGAACGAGGG  
 ACTCGGCCTCGGAGGCGACCCANACCACAGACACTGGGTCAAGGAGTAAGCAGAGGATAAAC  
 AACTGGAAGGAGAGCAAGCACAAAGTCATCATGGCTTCAGCGTCTGCTCGTGGAAACCAAGATAA

AGATGCCCATTTTCCACCACCAAGCAAGCAGAGCCTGTTGNTTTGTCCAAAATCANAACTGCACAT  
CCACANAGCANAGATCTCAAANATTATGCCANAATGTCNNGAATAAAGTTTCTGNAAGAGAGCTC  
TGCCTTTTTCTCTTGAAACATGCTTGNAC

SEQ ID NO: 926 GTACGCGGGGTTGAAAAATGGCGACTGTGGCAGAGTTGAAGGCTGTTTTAAA  
GGACACCTTGAAAAAAGGGGGTATTAGGGCATTTAAANCAAGGATCCGAGCTGAAGTTTTCA  
NTGCCCTAAATGATNACCTGATAACCCNTANCCCGNGGTCNCCATGGTAGGCACGGCNACTACCA  
TCAAAAGTTGATAGGGCAAACNTTCAAATGGGTCGTCCCCNCCCCCGCTACTTTTTTTTTNTTTT  
TTNTTTAGGGNCTTTCAATTTTTTATTTAAATGCCTTGATCCAAANATGGNTTTNANNNTTTGNCTAA  
NNNAGGAACATGCTTNGNCATGGAATGATCTTCNACTGCATAAGACAACATGGTNGAGGTCTGAG  
ATTTTACACAGAGTCATTCAAAAAGTTGCTTGTGGACTACTATGTGAAAGATAACAAGACACAAAT  
GTAACCTTCTTGAGGACAAAAAAGGATTTTTGTGTGGATGAGGGCTCGGCTACAGCAAANAGTANA  
GGAGCNGCTCAAAAAGAAATGTTTCACTTCTGCTCTGCTCTATGATCCCAATTCAAATGCNTGACC  
AGTGAACCCGTGNAGGCNCAAAAGGGTTTGAAAACTCNCAAANGTCCTGGTNGGTGAACAACN  
ACCTTGCCAAGATT

SEQ ID NO: 927 ACCTAGCTTCTGATGTATGCAAAACACTGCAGGAAGAGAGAATGAAAGAAA  
GAAAAGCTCTAGCTATAGAAGGAATTTTAAATCAAGGAGAAGAGCAAGGCCCCCGCTACTTTN  
TTTTTTTTTTTTTTTTTTTGGGAGTTGGANAGTATAACCTGACAGCATCTAACAAGTTTTGGGTG  
GGAGCTTTAATCTGAATAATCATATATTTTTAAACTAGGGACAATTTTCTTCTATTTTCAAAAAT  
ATTTTCTACCACCACCTATTTATTCCTATTTATTTGAAAGTGTGACCTCCTTAAATTTCTCTCTATA  
TCTCTTAATTTTTCTCCCATATTTTCCAAACTGTTCTGTATTGNGGGACATTTCTGGATTGGTAGG  
AAACTACTTGATCATCAGCTTGTCTTATATTCTGAACTCTATTATATGCCTCTGAACATCATGTTT  
ATCTCAGGTTATACCTACTTGATTGAAGGAATAAAAAAAGGAAAACCTGCTTCAGTTTCAAAA  
AGGATTTAACAACCCCTTTGAACAGAATTTTAAACATTTTCATTCAAATATTATGTTACTTAAACNC  
CNNTNAAAGAATTTAAAGAAAAATGGANACNATGCTCATGAATGACTTTGANAATGGATTTTAAAT  
C

SEQ ID NO: 928 ACTTTTTTTTTTTTTTTTTTTTTTNGGAACACAAGGGTCAGTTTCTTCAATTCATG  
ANCAGTCANAACAGGANATGCTTAGGAAGGAATCGTGGCTGGTGCCTCTTCTCCATGCTCATCCC  
ATACCCAGTGACAGGATACCGCTTCCCTGAAGTTAAAAACATGCACCACACTTCCGGTAAAGGC  
TGGAGCCACANAGGGCACCTANCAAGCTTGCTTTCAAGGTCCCATCAACAAATGAAAAGGGCAGC  
CGAGGACANATTGACANAAACCCCACTATGTTGCTTTTCTTTTTTTTTTTTTTCTTGAGGCAAANT  
CTTGCTCTGTTGCCAGGCTGGAGTGCATGGCACAATNTAACTCANTGCAACCTCCACCCCTG  
AGTTCAAGCAATTNTCTGCCTTGGCCTCCCGAGTANCTGGGATTACAGGCNCTGCCACCATGCC  
TGGCTAATTTTTNTATTTTNAGTANAANACGGANTTTAANCNTTTTGGCCANGCCTNATTTTNAAC  
TGTTACCTCAGGTNAACCACCTTCTTNAACCTCCCAAAATGCTGGGGATTACAGGCATAAACCCT  
NGCCCCNANCCCAAANTTTTAAAAAAGGTGAAAAGCTTTTNC

SEQ ID NO: 929 ACTGGTGCACCTCCTACATATCAAGGAAAAGCAAAACACAGAATAATTTAAT  
ATGCTGAATAAAATTTGTTTACACCAGATACTATACATCATACTGATGGTTGTCCAGTATGAATTT  
TAAGGGTATTATGTTAGATTCTGCAAAATATATTCCTATTATTCAAGTGAGGAGTCAAAGTCCA  
ACTATTCAAATGGCCATAAAACAAAAATGTTGCAGAAGGTATAGACCATTAAAAATAAAAAAGTCA  
GGAGTGGGCGAGCTGACCCCTTAGGAGCCTCAGGAATTCCTTTTAATGCAAGATAGATGGCAAGA  
GCTGGCTTTTTGGTTAAGTCAGCCAGTTCGGAAACCCATCAGGGAGAAGTTATCAGGTGTCAACTT  
GTAAGGCAGATGACATTCATCAAAAGCATCTTAAGAAGTANNANTGATCACAGAANGGAAAGCA  
TTCANGATTTNCACTANCNACAAATTGCATTAACCATGTCTAGGCCATTTGACAGNTGNTATTCTT  
AACGCTCCATANNCNTGAATCACANGGCCAAATGAGCCGACTAACTCTNACAT

SEQ ID NO: 930 ACTCGATGTGTAATGAAACCTGAAATAATAAGATAATAAGAAAAGCAATAAT  
TTTCTAAAGCTGTGCTGTCGGTGATACAGAGATGATACTCAAATTATAATAAACTCTTCATTTTG  
TGAATTATAGAAGCTACTTTTATAAAGCCATTTTTTTAGGGAACTAAGGAGTGACATAGAACT  
GATGAATGAGCAAAAAGTAAGTTTTGCTGGATTTTTGTAGAACTCTGGACGTTGAGGATTCATTATG  
CTGTGGTTAACTTTAAATATTTTTGAATTCCAAATATCTGAATTAATGAGCCTTGTCTTTACAAATA  
TGTGCCATTGTGCAACATCGGTGGATTTTCTAAAAATAATGTAAATGTCTTCTATTAAATGTTGAG  
TGCAATAAAATACAGAAGAATTCTCAANANATAAANTAATAANAAAAGTNCCTGCCNGGGC  
GGCCGTCGAAAGGGCG

SEQ ID NO: 931 ACATTAGAAAACACTTGTGACATTATTTCTAAGTGCAGGAGAGCAGCTCCT  
GGTGGGAGAGTAATGAAGTTGTTTGTATAGTGTATGCCAAGGATTTACAGCACTCTAGAATTTTC  
ACAACCTCTCCATGTTAGTGAATGACATTAGGTAAATGTTGTATTTGCCTACTCTCAGCTATCAGAT  
GTGGGATTTCAAACCTAGAAACAAATAAAGCATACACTCAACCAACAAGGGTAGTGCAAGTGCTAA

GTAAACAGAATGACATGGAATAAATTCATAGTTGCCAACTAGTGAATGCTTCAGCAGGGAGACCC  
TGATTTTATGATGTCTGTGACTTGCATCCCTGTTCTCAAATGAATACTGTCACTGCTCCATGTCAGC  
TACAATGACCGGAAAAGGCTGCTGGGGTGAGGGCTGANTGAATGAGGAAGATGACAGGAACTCA  
TTTTCCAAACAAGTGGTTCAATTTATCTGTCACTCTGTATATAGTGTAAAGTCTGATTTGTTAACA  
AGATCTCAAACCTCTCCANAAAAAGCAGGTC

SEQ ID NO: 932 ACGCGGGGGGAACTGCAATTGGTGGCTTTGAAGGCGCGGCGAGCGGGAAC  
AGCTCTTGAGGAGTGAGACTGCAGGAGATGTGGGCCGTGCCAAAGAGATGGATGAGACTGTTGCT  
GAGTTCATCAAGAGGACCATCTTGAAAATCCCCATGAATGAACTGACAACAATCCTGAAGGCCTG  
GGATTTTTTGTCTGAAAATCAACTGCAGACTGTAAATTTCCGACAGAGAAAGGAATCTGTAGTTCA  
GCACTTGATCCATCTGTGTGAGGAAAAGCGTGCAAGTATCAGTGATGCTGCCCTGTTAGACATCAT  
TTATATGCAATTTTCATCAGCACCAGAAAGTTTGGGATGTTTTTCAGATGAGTAAAGGACCAGGTGA  
AGATGTTGACCTTTTTTGATATGAAACAATTTAAAAATTTCGTTCAAGAAAATCTTCAGANAGCATT  
AAAAAATGTGACAGTCAGCTTCAGANAACTGAGGANAATGCAGTCTGGATTCTNAATTGCCTGGG  
GAACACANT

SEQ ID NO: 933 ACTTTTTTTTTTTTTTTTTTTTTTTTNGAACCCAGTTACAAGAAACAGGTCTGA  
CTTTCTTGCAAAGATTCTGCTTCCTCCTCAAAGTTCAATCTCTCTGCTGATCACCGCCATCTGCTGT  
TCTTCTATGGTGCCTAANAACCTTTCCAACTACCAAGCCCCCTAANACTGTGGTTTGATTCCANAA  
AGCCCAGGTGCCTGCCTGGCAACACACAGGCTTCGTCCTCTATTTTCCTTGGACATCATAAGG  
ATCATTTTGTGTCATCCCTCANAGCCTCCCTTAGGAGCGAATATGCANAATCTCTTGAATGICTT  
CTTGAAGTTCTCATTGCACAAGGGGTAAATGAGGGGGTTCAGCGTGGAGTTGATGTANCCACCC  
AGATGGTGAACATGTGCAAATGTTCAATTGCAACAGTTCTTGCAAGGCAATGACCATGAAGAAG  
ATGAAATAAGGGATCCANCAGAGGATGAAGGCTGCCATGATAAAACCCAACTGTTTGCGCGCCTT  
CCTTTGNGGTTTCATGTGCAACCCAAATACATACTGTCTTGAATGCGAGCGGAGCCTNTTCCAAGT  
AAAACCTGATGTAATCCAGGCCTGGGTANACCCACTCCTNANTTGGCNTTGGCTGGGGCTGNCTT  
TTGGG

SEQ ID NO: 934 ACCGAGCACTTTATTCAGTGCATAGCTTTAAGCCAGTGTGGATTCACTAAGT  
GGACAGTCAGTCTCCAGCTCTCTGCCTTCCCCAAAAGGGTCGTAGTAGGTCACCCTTCTACAGCA  
GCTAACTAGAGTCCTAACTAATGGGATCCAGCAGGGCCATTTCTCCAGAGGGCCAGTATCCTATTA  
GGAGACTCTTGGAATTCTTAGGTTCTACTCAAGAGTGGAAGGACCAATCACCTCTGATATTCTGTG  
GAAGGTTTTGGGGTCAAATCTGCCCTCTGCATCTGTGCAACTTGTATAAAAAGTCAAGTTAGTAT  
TACATGAATTTGGGGTAGGGTTAGTGCTTTGAAAAAATGTTGAACCGGCTGGGCGCGGNGGCTCA  
CGTCTGTAATCCCACTTTGGGAGGCCGAGGCGGGTGGATCATGAGGTGAGGAGTTCGAGACCA  
GCCTGGCCAAACATAGTGAAACCCCATCTGCTAAAGATATAAAAAAATTACCTGGCGTGGTGGC  
GCACGCCTGTAATCCCACTCTCGGGANGCTGANGCAGGANAAATTGCTTAACCT

SEQ ID NO: 935 ACCCTGGCATTGCTGACAGGATGCAGAAGGAGATCACAGCCCTGGTCCCCAG  
CACCATGAAGATCAAGATTATTGCTCCCCAGAGCGGAAGTACTTTTTTTTTTTTTTTTTTTTTT  
GGGGATAAACATTTTATTTCAATTTAAGGTAGTAGCAATACAAAATAAGTTTTGATAATTATAAAG  
TANAGACAATGAAAAATCCCACCCTATTTTANAATTTTAAAAATATTGAGCATCATTCTGTCTGA  
TGGGTTTTTATGTAAAGNNGGAATTTACATGACACCAGTATCTTGGTAATTCATTTTTTTGAATGC  
TACATATCCCAATCACCATATATTAATTTCTATTTTCTTCAAAGTGCTATTGCCACAAGAGACTG  
AACGATACAGCANTTTTAAAAAAGACCCACAACTAAATTGTAATGTCTTCAATATCTCAATTACT  
GTTTTTTAAAAATAAAAATGCANTAGTGAATTTCTTTAAGTTCTTGCTACATGATTGAATGGGAG  
ATAACCTCAACAATTTTTATTCTTAGGCATGANCAINTTCATGCTAAGGGNAGTCTATACTTAAAA  
GCNTTATTTTTTNATAAACCTGGGGCACAAATNTTTAAATCNTAAATAATGCNAAATAAAATCA

SEQ ID NO: 936 ACATGAGATCAAACCTGTATACAGAGCTGCCATAACGTATAAGTTAAGCCATG  
TATATACATTAATACATACATATATGTTCCAATAACATGTTGAAGATTTCCAGGTTTTTTTTTCTT  
TAAGTTATTGCTAGACATATCAAAACACAGTATAAACTGGTCATCGCAACACCCTTCTGCAGTAA  
TGATGAGAGTGGGCTAGGAATGTGATGAAAGGCACAGAATTCATAAGACTTGAGTAAGTAGATCC  
AAATTTGTTATCACTAATGGCTCAAACAATGTGGAGCCACTGATTTTCTGAAGTGAATATAAGAAA  
CAGTAGTTAATAGTATTTATCCAGGCTTCTAGTATAAGACATTGT

SEQ ID NO: 937 ACGCGGGGGAAAATCACTGTTTAGTCTTCTGGAGGCTATGATTTTTGCCTTAC  
TCCCAAAGCCACGGAAGAACGTTGCTGGTGAAATAGTCCTCATCACAGGTGCTGGAAGTGGACTC  
GGAAGGCTCTTAGCCTTGCACTTTGCCCGGCTGGGATCTGTCTTGTCTCTGGGATATCAATAAG  
GAGGGGAATGAGGAAACATGTAAGATGGCTCGGGAAGCTGGAGCCACAAGAGTGCACGCCTATA  
CCTGCGATTGCAAGCAAAAGGAAGGATGTATAGAGTAGCCCGACCAGGTTAAAAAAGAAAGTCG  
GCGATTTTCCATCCTAATCAACAATGCCGGAATCGTAACAGGCAAAAAGTTCTTGAAGTGTCCAGA

TGAGCTTATGGAAAAGTCATTTGATGTGAATTTCAAAGCACATTTATGGACTTATAAAGCCTTTCT  
ACCTGCTATGATTGCTAATGACCATGGACATTTGGTTTGCATTTCAAAGTTCAGCTGGATTAAGTG  
GAGTAAATGGGCTGGCAGATTACTGTGCAAAGTAAATTTGCAGCCTTTGGGTTTGTCTGAAATCTGT  
ATTTTGTAAAAACATTTGTCCAAAAACAAAAGGGGGATCAAAACCACGATTGNGTCCCCCTTTTIT  
TATAAAACTGGA

SEQ ID NO: 938 CGCGGCGAGGTACCCAAGCCAGAGCTGAGACATGGCTCCCCAGATGGACTG  
GCTGTGGTCAGGAAAGGCCTGTAGAGGGAGCTGAGGGCTCAGAAAATACCTGCTCCGGGTGCCTG  
GGCTCAAGTTCTCATTTCTTTGATGCCACTGGCCACTGTATCTGCTTTGTAAAAACCTCCA  
AAAAGTTATACATCAGGTATCTCTAGTCANCTTCTCCGCTGCCACCTCTTTGCCACAGATGAGTG  
ATCACTGTGAATGGTGGTGACCAGGTTGGGCGANGGCAGGGGCTNGTCTAAGTGCCTTAACCCAG  
GCCCCNGCCCATNANACCTTNAACAGNAATTTTNCNTNCTNTTAATNCATTTCCAATTCAGTTT  
AACTGGGGGGTANTGTTNATACANAATAGCTNGATTAGAAAAGGAA

SEQ ID NO: 939 CGCGGCGAGGTACAGTGACCTGCAGAACTTAGCCAAGAGTCTGGGTCTCCGG  
GCCAACCTGAGGGCAACCAAGTTGTTAAAAGCCTTGAAAGGCTACATTAAACATGAGGCNNGAAA  
AGGAAATGAGAATCAGGATGAAAGTCAAACCTTCTGCATCCTCTTGTGATGAGACTGAGATACAGA  
TCAGCAACCAGNAANAAGCTGAGAGACAAGCCACTTGGCCATGTCACCAAAACAAGGANAAAGGT  
GCNAGACTGTCCGTGTGGACCCTGACTACAGAATCATGATAAGCAGGAAAGCCAGGATCTNAAA  
GCTTCTGCAAAAGTTCCCTTCTCCACCATACGAGCCCCAAGAANCTGTNAATGCTNTTCCCTAAGT

SEQ ID NO: 940 ACTGTCTCTCCCCAGAAGGCCTTCAAGGTTAACACACAACANTGCCCTGCC  
CTTGATTACTGGCTGGGCTTACAAAGGCGAATACTCAGACAGAATCAGGAGGCACTCACGCTTA  
ATGAAGGGGCAGGGATCTTGACGCCAGCAGATTGCCTGGCCTATATTCCTNTCCCTNGGACTCTT  
CTGAATTGTCTTTTGGGTGTGGNTAGCTGCNAATNNCANCATACATTNCTTTATATTAGTAATGA  
AANATNGATTNCTTATTGCA

SEQ ID NO: 941 ACTCATCACTCTGTCCATACGCGATCACAATATCCTCTAGTTCTTCCATCACA  
GTCTGCGCACATTTGGTCATCAGCTGGAGAGCACGGNTGTCAATTGGGTTTGTCAAAGTTGTGCTTC  
TCATCAAACCNATGGAAATTCGGGCCGTCCAGCCGCACTACCACCCAGCNGTGTGCCAGGGCAGG  
TGTNGTNAGCNTCTAAGCCCTNACNTACCTCGGCCNGTGACCACTCTANGGGCNAT

SEQ ID NO: 942 ACGCGGGGACAGCGCGCGGAAGAAAAACCAGCAAGAAGGCGGCGGGGAA  
GATGGCGGTCTCTGGGGTAGAGTTTGAAGCTTTCTGACTAGGCTAGTCGAGTAACATTCGGGTCA  
TGGCGTCAAACTCAACTAAGTNTTTCCTGGCATATGCCGGCTATGGCNAACAGGAAGTGGATGCC  
AACTCTGCCCTTATGGAATTGGACAAAGGCCTNAGATCTGNCAAACCTTGGCNAACAGTGTGAAGC  
ATTTGTTTCGCTTTCCAGACTTTTTCAGAAGTATCCNTTCCCTATTCTTATCAATTCTGCATTCCTAA  
AGTTAGCTGATGTTNCAGANTTGGAAATAATTTCTNAGGCTATGTGTTCTTAAATTACCCAACAA  
GTTGATAAACNTTGGANGANATTNTAAGNTGATTAATTTGTGAANACAATTTNTNTGCATTTCATAT  
AATGNCNTGGGGNAANACCATTTCCTCGATTTTGNCTCTGGATNATTTTCTTNAATGAAATNTTTT  
TTAATAT

SEQ ID NO: 943 TTAAAAAAGTTCGAAAGTACAGACACAGAATTTAGGAAGCTGAAGGCTGAG  
AGTCTCCCTTCTCACTTAATCCATGCTTTATTTTGCATTCTCACAGGTAAGGAGGCAGTGCCTGTT  
ATGCTGTNGACCAANACCAGCCCCACGGAGCTGATCTTNAAAAAATGGAATTTACTCTGGCATA  
CTCCTATGTATGATACCTTTCCAANGNCAAATCCCAATAGACCTGNANGTGCAANTTTGGGCAATG  
ATCCA

SEQ ID NO: 944 ACCTCTTCTGTTTGAATGGGTTATCCAGTAAAAAAGGGCGTGCCCATGCAA  
AAGGAGGGGAAATCTAGAACTTTTAAAGATTCCCAATTTTCTGCATTTGACTCCTGTAGCAATTA  
AAGCACTGTGAAGCCCTTAAAGATTTTNGCACTTGAGTGGGCCAGCCGCACTGGACAAGTGAAG  
AAGAAATGTGAAAAAGCATTTTCCAATTTGAAATTTGACAGCACTTGATTATGTTTCATCAAGAACC  
ATCTGTTTCGGAACCCAGAACNCNAATTTTATTCTNTTAAANTAAANCTTTCCAGTTTGAATTTNN  
ATTNANCAACNCNAANAAAAAATTAATTTAACTTNTAGGAANAACCAATACTTCAAAACCCAA  
NATTGTGCTTACCATCNAAANAAANANGTCCCTTTAAGGAGGGGCNAATTAACCATATNCCAGN  
GTNNCTACTAACNNNNNTTTTACCATNAGTCTTTGGAATNCTGAAAAAATGGGNAAAAAANTAA  
GACTNAAGCCNGACATTGGGAANAATTTNTTATGGGGAAAAATNGCTTATCNCAAAANAAAAATAT  
TCTTGGAAACNCTTCTTCCANATTAAGAGCTNCTTGANAAAAATATGNGAAATTAATAAAGANA  
AGCTCCT

SEQ ID NO: 945 ACTGGAACTAAATCATATTTCTTCCCTCCAAATTTCACCCATTCCTGACTTTG  
AATCAATTGCAGAAATGCAGGTGTGTTACTTTGTTGATCAATAACTTTGGAACAATTATGGATCAA

TTCTATGGTCACTCTGAATTTTCATGTCAATTAATCACATAAAAAATTGATAATACCTCATTCTGTATT  
 ACAATATGATTTTATTTTGGCAAAGGCAAGACACCTATAGTTGAGCTGTATTTTGGGGGATTGGGT  
 GAGGAAGGACTTCTATCTTATCTCAACAAAAAACTGGCCAGTATTTTGTAAATGTAAAGCTTCCT  
 TTTCTTTCTAAAAAATAGTAACAAAAATTATTTTTCATTGGCCTATTCTGTTCTTGTGTCTAAACTAA  
 CATTACATTAATTTTAAATCTTAGTTTCTGATAAACACAAGCCATTCTATCAAAAAATATTATTTAT  
 TTCAGTCAAATTTTACCAAANTAACAAAAGACCATTATNTCCGTTTTTTTTTATTANTGAANCAT  
 NATGATTTTTTGACANGCTTGTTCCTCGTCCGTATAAAATTTTTTCCAATCAAACCTACTTTTTTC  
 CNTACTCTGGGCATATTTTTTTGNGAAGGTTATACACATTGAANAACCCTTAAAAATCCCN

SEQ ID NO: 946      ACGCGGGGAAATCAGCAAACCTGGGAATTTCTGGTGACATAGACCTCACCAGT  
 GCTTCATATACCATGATATAATCTGAAAGGGGCAGATTAAAAAAGAACTCTAAACCTTACAT  
 GTGTAAAGGTTTCATGTTCACTGTGAGTGAAATTTTTACATTTCATCAATATCCCTCTGTGAAGTCA  
 TCTACTTAATAAATATTACAGTGAAAAAANTANAATAAGAAAATAN  
 AACAC

SEQ ID NO: 947      ACTAAGGGGACAATACACCAAATTTGTTGAGTTTACAATCAAGTCTACTAAG  
 GTTGAGACTTCCTTATCAGTTTGGCGAGTCCCAGGGCAGAATAATCATCCATCTACAGGTCTCTGTT  
 TCCTCTCCCTCCACAGCAGTGGAGAGCATCCCAGTGTTTGGGGCACTGTGTTCTCTTCGTCCCTGC  
 ACCAGACCCTGGAAGCCTTGGCCAGAGACCTCACCAAACCTCGACTTGC GGCGCTGGGCCAGCTTC  
 ATGGATGCTGGAGTGGAGCAGATGACGTAGCAGAGCTGCTGCAGGAGCTACAAAGCCTGGCCCA  
 GTGCTACCAGGGTGGTGACAGCCTCGTGGACTAAAGTTCCCAAGTGTGGGAGAAAGGAGCTAGTTT  
 GCAATAAAAAACAGCTGGATGCAGGAGCCANTGTCTTCATGCAGAGGAGCTCAATGTCGCGGGAC  
 TAGCTACACCAACATATGCACTTTTTACATTTAGAAAACACTGTGATTAGACCACAGAACAAATTAA  
 TATGTGCCATCANACCAAAAAAAGTNNAGAAAGGGAGCTGAACTCCNTCTTCGATGCTATTNN  
 CAGAAGGACATTNTGTAAAAGTNNNNATAAAAGACCTTGNATTGATGCCTTAGGNTGGCAANAA  
 CCCCTGGG

SEQ ID NO: 948      ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGATATTACACCATAGGTTTTATTA  
 CGATAAATGTTTGCATTACTTTTAAAAGCTTAGCTCTTACTAAGCATTCTTTAAACAAAAGCTAATA  
 ANCAAGAAATCATTTGCCATACGAAACTATATTACAAACAAGACTTTAATCCAATATTGAAAG  
 CTAAGAATTAGAAAAAATACAAAACACTGCTATGAGTCAATTGAACTGCTATCATTGAATTTGCT  
 GCATTTAGAATGACATAAACATACTGAACATAAAAAACAATTTATGGATTTATTCTATAAGACTAG  
 CATTAAGAATGACATACAATTNGTGATTTCCCTTAAAAATAATTTTTTACAACANAATCCATTTGA  
 ACAAGGGTCTTTTTTTTTCCCTCATTTGAGGGGAAGACAATCTATGTTTCCCAAACAGATCCTCC  
 TTTCACTAAAAATAGCAAACCTGNGGCCTCATCTCCTNTCCAGATGCTACTTATAGATGACTTT  
 GCATAATAAATAAATAAGAATTACTTTTTCTGGTAACAGTGTCAACGGCCATAAATAATCAGTTT  
 TAAAAAACAAACATCAAAGTGCCAAATNTAAAAAACTTCCTTTAAAGAATTACCCAACCCCA  
 G

SEQ ID NO: 949      ACACATCAAGTCAGAATGGGCTAGCCCATCAGGGAAGCAGCGGTAGAAGAA  
 ATCTGGGCGTGGCCTCCCTACGATCAGTTTTATTGTGTTGGTAAAGACGCCATTACAGAGCCAGGGC  
 AAGGCTGGCAGCCAGGCAGGCTTGCTGCTGTCTTGTGTCTGCCTTAAACATCGGCTTGGTGGG  
 GAAATACTCCGCTCCACGTAGGGGTTCGGTAGAGCCACATCTCCTCCGGCTGGATGAGTCTCTG  
 GAACGGGGGAGCAGCTCCGTACCAAGGAAGGCCGCGAACAGCGCGAGCCGCACGCCACCCCC  
 GCGT

SEQ ID NO: 950      ACTTCCAGCCAACCTCGTGAGCCAGGCGCCCCAGATAGGCAAACCTTTCTTGT  
 AGGCTTCAGACGCACGACCTTGAGGGCAGCAGGAACCAACCATCCGCTTTTTCTCGCACGGTCCGC  
 CAGAAGATGCGGCTGGGGGCCCGAAGTGGTAGGGGCCCTCGGAAGGGTTGGTGTTCATCCGCTT  
 GCGGANAAANCCAAGTNCCTTGGCCNNACCACGNT

SEQ ID NO: 951      ACGCGGGGGCCTTACAGTTGCTGAGAGGAGGCGAGAGGCGGGGGCGCTAGG  
 GCCGAGATCATGTGACTGGGAGAGGTTTCCTTGGCAGCAGAGGACGCTAGGTTTGGGATGAAA  
 GAAGCTGGGCGAGATGCAAAATCTGGAGAGCGCGAGGGCCGGCGGTGAGTCAGCACCCAGACTG  
 GCAGCATGACCGGTGAGATACCAAGGCTTTCTAAAGTCAACCTTTTCACTCTGCTCAGCCTCTGGA  
 TGGAGCTCTTTCCAGCANAAGCCCAGCGGCAAAAAATCTCAGAAAAATGAAGAAGGGAAAAGCAT  
 GGACCCTTAGGAGATAATGAAAGAGAAGACCCAGAGTNTCTACTGACAAAAAGACANGTAAAGA  
 AAAACTGGTCTTGTGGTGGGTGAAAAACATTAATAATTGNTGGGTNTCNACTGTTCTNAGTGAAAA  
 TTTACATTCNNGGGCCAAATTGCTTTTTATNAAACATTGGGTAAAGGGCTGAAAAACTGTNANCTT  
 TANTTTTTCANAAAACCATGTTTGTCTTGTGTTGAAAAATGATTGGAAAATGGATT

SEQ ID NO: 952     ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATGGCCCANGCTGCCTTCCT  
GTCCTCCCATCCTGCATCCTGCTTTCTGTTCCCTGGATACCGTAGGATGGTTTTATTTCAGTTTCATG  
CNCAAATTANTCTGGACACTGNGGAGTCATAACAANAGTGGGATGGAGGTTCCAGGGCANTCATT  
TTCTTTGGAGGGAAGCTTNAAAAANTANGTTAAANCAATGCCCANAAAAANCCACTGNTTNTNATTC  
CCATCCCCAGCAATGANCTAAAAACCAACCTCAGCCAAAAAGAGGGTCTTTGATCACAGAGGGTC  
ACNNAGTGGGATGTGATTAGTGGGTTTGGGGATGGAGNCCANTTAAACCAAANGGCCAAAGGGA  
TCCNNNTANTNCNGGCTTAANGGGAATGGGGTTNAAAATTAAGNGGTTTGGGGGACTTAGGGCNN  
GGGGANCCCCAANGNNGAGGAAAACTAAAAANGNTGNNNTGGCNAAANNTACNTTGGGCATTTA  
ANCCTGCNTGCTGNTACNNCTANCCCCCATTTATGCCTTTGNTAAAT

SEQ ID NO: 953     ACAGTTGAAAGCAGAGTGTAACAAGGGATATGTCAAGGTAAAGCAGGTAGG  
AGTCAATCCCACCAGCATTGACTCAGTCGTAATTGGGAAGGACCAAGAGGTGAAGCTGCAGCCTG  
GCCAGGTTCTCCACATGGTGAATGAACTTTATCCATATATTGTAGAGTTTGAGGAAGAGGCAAAG  
AACCCTGGCCTGGAACACACAGGAAGAGAAAGAGATCAGGCAACAGTGATTCTATAGAAAAGG  
ATGCTGCTCAGGAAGCTGAGGCTGGGACAGGGCTGGAACCTGGGAGCAACTCTGGCCAATGCTCT  
GTGCCCTAAAGAAGGGAAAAGATGCACCTATCAAAAAGGAATCCCTGGGCCACTGGAGTCAAN  
GCTTGAAGATTTCTATGCAGGACCCCAAAATGCAGGTTTACAAAGATGAGCAGGTGGTGGTGATA  
AANGATAAATACCCAAAGGCCCNNTACCATTGGCTGGTCTTACCGTGGACCTCCATTTCCAGTCT  
GAAGGCTGTGGCCNNGGANACCTTGAACCTTAAAGCATATGCACACTTGTGNGGGGAAAANGTGA  
NTGGTANATTTTNTCTGGGTNCAANAACTCCGNTTCCAATTGGCCTNCCGCCTT

SEQ ID NO: 954     ACCTAATGAAAAGATCTCCAAGAGGTTTGCTCATTCTCCTTGGGCTGTAAAA  
AAGATTAATCCTATATGTAATGATCATTATCGAAGTGTGTATCAAAAGAGACTAATGGATGAAGC  
TAAGATTTTGAAGAGCCTTCATCATCCAAACATTGTTGGTTATCGTGCTTTTACTGAAGCCAATGA  
TGGCAGTCTGTGCTTGCTATGGAATATGGAGGTGAAAAGTCTCTAAATGACTTAATAGAAGAAC  
GATATAAAGCCAGCCAAGATCCTTTTCCAGCAGCCCTAATTTTAAAAGTTGCTTTGAATATGGCAN  
GAGGGTTAAAGTATCTGCACCAAGAAAAGAACTGCTTCATGGAGACATAAAGTCTNCNAATGTT  
GTAATTAAGGCGATCTTGAAACAATTAATCTGTGTATGTANGAGTCTCTCTACCNCTGGATGAA  
AATATGACTGNGACCTGACCNTGNGGCTTGTTACATTGGCCCANACCCATGGGAAACCCAAANAA  
GCNGTGGANGANAATGGTGTATTCTGACNAGGCNNACATATTTGCCTTTGGNCTTNCCTTTNGGG  
NAAAATGATGAACCTTTTNTCNTTCCCACACATTTAANNNTACAGATGATGATGNATGATGAAANA  
AACTTTTGNNGAANA

SEQ ID NO: 955     ACAACTCTTGCTAATGGAATGCTATAATGCACAAGGTCAAGGATTTAATAAA  
TTCTAAAAGTGTCTACATATATCAGTGATAACTGTATTATTAGAAATATAAATGTATAGAAATATA  
AAGTATATGGTATTAAAAACAGACCTTGCTAATATAAACATATATAAAGTATGTCACCTCTCCTGT  
AATAACAGCATAAAGATCGATCTACAGTTTGCCCTTCGCTGGCACTCTTAAACCACTCCTCCAAT  
GGTCAATGTTGACCTTGAATCAACAGCCGCTGAACCCAGGAGACCCCCACAGATGTGTAGATTCA  
GCACCTAGAGGGCCCCCTACCCTCTTGCTGTGTGTTCCCATGACTCCAGAAATAATTAATCGC  
AACTTGCAATTTTAANGTCCACAGGCAAGTTTGAAATCTAACCTANAAAAAGTGNANGCANAGGCA  
AAATACGCGGGAATTTGTTATAAAAGCAACAAGATTTNCTTAAAAATGCTTCCAGTTCAAAGTCAA  
AATTAAGGTGACATNAAGGTCCACCANCTTTTACAGAAAGTTGGGGATGTTTTGNTGNTTNNCTTT  
NAAAAAGAAANAATCTNCAATAAACATGTNNATTTGAAAAAAAATCNTGNGNTAACTTTTTTG  
ACCCATCCCCAA

SEQ ID NO: 956     TGCTGAAGCTTCACAGGGCGGCCAACTAACTCGCTGATTTTTGCAAGACCA  
CAGTGTAAGGTTCGGATGTCCACCTGAAGAAGGGGTGGGTGCAACTCTCTGGGTGCTGCACACAC  
CATGACCANCTTGGGCATGCATCACCCAGCTCCCATCCATTCACACTGGTTGCCTTNGTGAGGTC  
CATTTNAGAGGGCTTTCATAGGCCCTTTTAAATGAAAAAAAATATCTGGTCTACNACGAACCT  
AAGATNCATACACCTCCTATTTAT

SEQ ID NO: 957     ACCAGCACATGAAGCCCTTCTACAAAATTCCTGACGGACTGGGAATAAAAAAT  
TCCTAGTGACAGCCACTCCTTCTCAGGCAGGTGTGATTGTTTGAAATCTCTCCCAATATTGAGAT  
GAAACCTGCTTCCCTGTAACCTCCCTGTAATTCTGTGGGTCCCTTGAGCCACAGAGAAGGCAGCA  
ATCAGTAGGGGAAGTGCTATAAAAAATATACTATCCCGGCCAAGCGTGGTGGCTCATGTCTGTAAT  
CCCAGCACTTTGGGAGGCCAAGGCGGGCAGATCACTTTAGGTCAGGAGTTGGAGACCAGCCTGGC  
CAACATAATGAAACCCCGTCTCGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG  
GCTGANGTGGGTGGATCACAAGGTCAAGAGATGGAGACAATCCTGGCCAACATGGTGAAACCCC  
ATCTNTNTAAAAATACAAAAATTAGTTTGGGCGTGGTGGCGTGTGCCTGTTAANCCCANCTTCNTN  
GGANGCTGAAGCATGAGAATTGCTTGAACCTGNGAGGCAGANGGTTGCAANNGANCCCATNATC  
ACACCANCTTGAANTTACCTGGGT



SEQ ID NO: 958 ACCACAAAGGGACCCAAATTCAGCGGTCTGTGCCTACAACTTCATTAATAA  
CTGCTTGCAGATTGGCAGCTATCTGGTCACTTGACATATCCAATGTTGCTATTTTGGTCTGGAGAA  
AGTTCTCCCTTTCTTCATCTACCTTAATTTTCATGTCCATTTTAAATAATTCAAGCATTTTGGGGATG  
TCACGGCCAATGGAATCTTGAAAATTTAATTCATGCACTGAAAAATGTGAATTATTTAAAGGGTA  
AGAAATAGAAATGAAGACTGTGCCACTGAAGAGACATCACTCCCACCATGCTCAAGCAAGTCTTCC  
ATANATGAATCAATGTGAGGAACTCATTACATAGAGAGCACGGCATACCTCAGCATACATTATT  
CCTAGGATTATCTCAGCAGTTGTTCACTTGGCTCANGAGCACTACAGCTATAAGTGAAAGCCAACT  
ACCCATAGGCCAGTTTAAATTTTCNAGAAAAGCTTTGGATATTTTTTATTTCANTTCTTCCTTAAT  
CGATTAAAGTTTCAGGCATTATTTCTGGAACAGCTACGTAAAAAGTCTGCAACAATTCATCATCCC  
AAATCTTCTGTATCAGAACTAGTGCCTCCTGCNAATGCANCTTCATTTTCTCCGCTATTTTACCT

SEQ ID NO: 959 ACTGGAGCTATACTGACATCAGTAATAGTGACACCCCTGGATGTTGTTAAAA  
TTAGACTCCAAGCCCAAAACAACCCACTCCCCAAAGGAAAATGTTTGTATATAGTAATGGACTC  
ATGGATCATCTATGTGTCTGTGAAGAGGGAGGCAACAACTATGGTATAAGAAGCCAGGAAATTT  
CCAGGGAACATTGAGTGATGGCAGTTCCTGCCACAGTTATTTATTTTACCTGCTATGATCAATTAA  
GTGCTCTTCTGAGATCTAAGTTACGAGAAAATGAAACCTGCATACCAATTGTTGCCTGGAATTGTA  
NCCANATTTGGTGCANTNACTGTGATAAGTCCACTAGAATTGATTAGAACCNACATGCNATCCNN  
AGAANTTTTTCTTACNTTGGAACTGCATNNGATTTGTCAGCATGAAAAGTATCTTGANNATGGTNG  
GATTTCCCTTTGGGAAGNNGCCTGGGCTCNTACTGTTTCTTATAGATGTNCCTCNGGCCGNGACCA  
CTCCTAAGGGCGANTTCCAACACNCTGGCTGNCGTTACTAGNGGATCCCACTNCGGTNCNANCTTT  
GGAGNAATCATGGGGACAT

SEQ ID NO: 960 ACGCGGGATTAAAAATTTCTTGATTTTCTTGTGCATTAATCTGACGATAATTT  
CCCTGTATATTATGTTCAATTTAGCTGTTTGTAATTTTGTTAATTAGATCAGGTGTCTGCATTTGTT  
GGTGTAAAGTGAACATCATCACAGTTATCCTGAGTTGAGTTTAAGCCAAATACATGCATAGAAAAG  
GGTCTTCCTATTAATGGAAGAAGGTAATNTTAGGATGTGTATTATTTTCAGTTTTGTATGTTTAACT  
TTTATTAAATAAAGTGTTTTTAAACCANAT

SEQ ID NO: 961 ACAAATCCAGCAAAACTGGAACAGAACAGGGTAAGAAAGCACATAGCAACT  
GCTTTTTCACAGTAAATTTTGAATTTCCCTTAAATATCTACCTAGTGGCTTAAAACTAAAACTAT  
GGTCTTTGGCTTTAAAAAGAAAGTCAAAAAATTTTGAATTTTAAATGACTGCACAATTCTACCAAAG  
CCAGTGTTTACCTGGGTATTTTTTGTACCTGTAGTTTACATTCCCTGCTACTGTTAAAGAAACAGA  
ATTCTACAGTATNCAATTCTGTATATTGTCTTTTAAAGGTTTTCATCANGACTCACTACTACTGCT  
TGAGGAGTCCGTTGACATAATCTCTACATCATCTTCATTTTCTTTATTATGCCAGGAGGTTCCAAC  
AAAAAGTCACTACTATGATTTGGTGGTAACATAATTCATCGACATGTCATTTTGAAGTCCCAATGG  
TTACTGTGGGAGCCAAAGGACNTCTGGCAATCTTCTGCTGCAANGTGCNTNTTCTGGTAAATGG  
CNATTCACGCCNTTATAGTGA

SEQ ID NO: 962 ACGCGGGATATACACTGGAACACATGCATGCTTTTGGAAATGTATAATTACCT  
GCACTGTGATTCAAGACAGTGTCTACTATATTGATACCTTGGAAAGAATTATGAATTT  
AACAGTAATGCTGGACACTGCCTTAGGAAAACCACGAGAGGTGTTTCGACTTCCTACAGATTTGA  
CAGCATGTGACAACCGTCTTGTGCATCTATCCATTTCTCATCTTCTACCTGGGTTACCTTGTGAGA  
TGGAAGTGAAGATTGTATGTCAATTGGAACAGGTGAACNGTGGAAATAGCGCTTCTGAAAAATGG  
GAGATTATGTTTAAATGAAGAACTTGGGGTCTTTTATTATAATTCACAGTATCTCACTGCTAAAT  
GCTGAAGAACATTCTATAGCTACCCTACTTCTCGAATAGAGAAAAGGGAATTGGATATGAAAGG  
AAGTGGTTTCTATGGTTCCTCTGGAGGGGGTCACTATCAGGNNANAAAAATCAAGATAATNAAA  
AAATTTGAAAATNTTAANCNGTGATATTTNTNCGTTGGAAAGTCAAGTGCCACATTNTNGCTGCT  
ATTGGACCCTGATGGAAGTGGTCTAATTGANTTGNNTNCNTACAGNGTCTTTNACATTTTGTCA  
AG

SEQ ID NO: 963 ACTTTGACTTACTAGGGTGATTCAAAGTTTCAGGAAAAAGAAAATTTCCAGT  
ATCNTTTNCTTAATCTTATTAACCCCAATAAGAATGCCAAAAAATACAGAGCTCACATTTTGC  
TGGCATAACATTTCCAAATTTTAAATGCCTCCCTGACAGGTGAATTTTAAAGGATAAAAAAGCANAC  
NCTTNCAAAACATTCCTTGTGATGAANNANAAAAAGCCCTGGATAANTGGCCAGCTNCACTGGAT  
TTTTGTCTAAATTTTNCATTAACNTTTCAGTCTCA

SEQ ID NO: 964 ACAGTTGAGGAACTCAGTAACCAGATATTATCTGCACGGAGTTGGTTGCAAC  
AGGAACAAGAACGGATAGAAAAAGAGCTTTTACAGAAAATTGATCAGCTTTCCTTGATTGTTAAG  
GAAAACAGTGGAGCCAGTGAAAGGGATATGGAGAAGAAGCTCAGCCAGATGTCAGCCAGGCTTG  
ACAAAATAGAAGAGGGTCAAAAAGAAGACTTTTGATGGTCAGAGAACAAGGCAAGAAGAGGAGAA  
GATGCACGGGCGAATCACCAAGCTGGAGTTACAGATGAACCAGAACATCAAGGAAATGAAAGCA  
GAAGTTAATGCTGGGTTTACAGCCGTCTATGAAAGCATAGGATCCCTNANGCAAGTTCTCGAGGC



CAAGATGAAGCTGGCNGGGACCAGCTACANAAGCAAATCCANCTGATGCANAATCCANANCCCC  
CATGTGAANGGAGCTGGGACAAGGTCCCTAAAAGACANGTTTTGCCNGTAGGGGCNTAGGAGCC  
GGGTACNCTNTGTTGCCAAGGCCTTGNTTGCATTNAGGATTGTCCATCCATNGGGTGCATNGTGCC  
AANAATNTGTTTNTATGGGNCCTAAATGNTTACCNTNGGGATNTTNGNAAATTCTCTTCTGTAA  
NNNTATNAAAAATCT

SEQ ID NO: 965 ACCTATTTCTAAACAATGATTNAAAGTCTNTATCCCCCTAAGCGGAGNTGTTGT  
NNNTCTCCCTAATCTATCACCTGCACTACTTGAGAAAATNTAAAGTGTTCNTANTTAAATNATTTT  
CTTCTTGAGCGATCTAATGTTTCTTGTAATATTGATGANCCCTACTAATNATCCTGCTGTNCTNTAAT  
TAACGCTTAATGAATAAAATGGCNCCT

SEQ ID NO: 966 ACGCGGGGACTTGTAAGGAGGAGAGAAGTCAGCCTGGCAGAGAGACTCTGA  
AATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTG  
AAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCTCTTGGCCTCCAACCTTGTTGGGCTACATCCTAG  
GCCTTCTGGGGCTTTTGGGCACACTGGTTGCCATGTCTGCTCCCCAGCTGGAAAACAAGTTCTTATG  
TCGGTGCCAGCATTGTGACAGCAGTTGGCTTCTCCAANNNGCCTNTGGATGGAATGTGCCACACA  
CANNCAGGCATCACCCAGTGTGACATCTATAGCACCCCTTTTGGGCCTGCCCCGCTGACATCCAGGC  
TGCCAAGGCCATGATGGTGACATCCAATGCAATCTCCTCCCTGGCCTGCATTATCTCTGGTGGTGG  
GCATGAATATGCACAGTNTTCNNCCANGAATCCCNAGCCAAAAACAGANTGGTCGGNAGCAGG  
TGGAGTTCTTTTTTCACTTCTTGGGANGGCCTCCTGAGATTCACTTNTNCCCTGAAAAATCT

SEQ ID NO: 967 ACAGCGGGATTAAAAATTTCTTGATTNCTTGTGCATTAATCTGACGATAATT  
TCCCTGTATATTATAGTTCACTTANCTGTTTGTAAATTNTNGTTAATTAGATCAGGNTGTCTGNATNN  
GTTGGTGTAAAGTGAACATNATCACANTTATCCTGACTTNAGTTTAAGCCAAATACATGCATANAAA  
AGGGTCTTCTTATTAATGGA

SEQ ID NO: 968 GGTACTTAAAGATGGGATGGAGTTGTAAAGTGCTTTTATAATACAATATAATT  
GTTAAAGGCAAGGGTTGACTCTTTGTTTTATTTTGACATGGCATGTCCTGAAATAAATATTGATT  
AATATGGCAAAAAA

SEQ ID NO: 969 ACTTTTTTTTTTTTTTTTTTTTTTGTCTGTCAAAACGTTTATTGCAAAATGGAGTCT  
TANAACAAAGGAAAGCAAAGAAAAGTTACATCAAAATGAAATGTATGACACCAACTTGGATTTC  
TGAATACACTGNGGACTTTGTGCTGGAATATCAAATTCCAACTATAAAATCAGTAACTGAAATAGT  
CTTACCACACAAGAGTAAAATTTAATCTTCCATACAAACATTATACAAGATATTTGGCATAGGACT  
TGCTCAGAATAACATTGCAATAGAATAATTGAGAAAAAGTTTTTGTAAAAAAAACAATAAAGNGA  
TAATGCCATNTATTCAAAGCTCACACTCAAAGAATATAAATATTTTCTATTATTAGTACC

SEQ ID NO: 970 GGTACCCAGTTGTTGTAGTGGGGACTATGATACTGTAATAATATTTTTTAAA  
AATTTACATCAAGAGAGGCAGTCATTCACGATGGTTTTGTGCCAGCTCTTTTAGGGTTTTGGATC  
ACATTANAGATATTTAGAACATATTACCTGTGACTTACGTAGGAAACCTAATATGCTGAGTATCT  
GGCACTTGAATTCCTGCTTTTATTGCTGGAGGTCCACATCTGTGGTTGACCTCTGTTATTGTTGAAA  
AAAANNAAAATGAGGANAAGGTTCTCCTTCTNCGATGCCTGTGAGGAGCCACCAACATTTGANGC  
TATGGANCTCATTGGTAAACCNAACCTACNNTGAGATTGGNGAACNAGTCTNTTATAAGTGTG  
AAAAAGGATACTNCTATATACCTACTATTGCCNCCCTCTATTTGNGATCGGAANCATACATGGCT  
ACCTNTTTANANGACCCCTGCAATATAGAAACATGTGCATATNTACGGGNTCT

SEQ ID NO: 971 ACAACGCTCACCTGAACATGAATTAGCTCTGACACATCATAAACCAAAATTA  
GGTCAATACAAGCTTATCTAAAGAGATATTAATACTTTCTTCAATTAAACATGCTTGGAGGAATAG  
ATGATAGAAAAAATCATGCTTCCTAAAATGCTGTATATGAGAAAGTAACTTGGCTTTTAAA  
GAAAAATATCTCTTTAAATACATGGTTTAAATAAATTACATTCCAATTGACTCAACATTTGGTTGA  
CTGCACTGAAATTGTGTAATTACTATGTTTTTGTCCCAGAAATACTTCAGTTTGCTCAAGAAAAAT  
ACACTCTCATGGAAAGTTTTTCAAAGTTGATGTCCTCCACATNTTCAAGTCACATTAAGGCTGAAC  
AAGCGTGACATATTTCCCAAATTATCACCATAACAGTTTATTTTCCAAACGTATTTTCTTAAAA  
AAAAAAAATNCACGCAATAGTTGTCTCTAAGTCTTGTTCCTAGCTC

SEQ ID NO: 972 GGTACAAGCAACAAATAAAAAATAGATAAATTGGAATTTATGGAAGTTAGA  
AAATTTTGTGCATCCAAAGACATGATCAAAAAGTGAAACTCAGGACAATGGCAGAAAAATATTTG  
CAAATCCTATGTATAAGAGTCTAGTATGCAAAATATATAAAGAACACTGAAAACAACAAAAACCA  
AACAGCACAATTCAAAAAATGGGCAAAGGACCTAAATAGAAATTTTCCAAAGAAGACCTACAAATT  
GCGAAGAAGCACATGAAAAGATGCTTGACGTTACTAGTCAACCAGAGAAATAAAAAATCATGAGAT  
ATGACTTTATACTAACCAGAATAAAAAATGTAAGAGAATAACAAGTGTGGTGAGGCTGTGGAGAA  
ATTGGAACCTTGTACATTGAAACAACACCTCTTACGATAAAAGGAAAAAAGTGTGAAAAATATGTG

TTCCAGGAAAGAAAAAGCACAGAACAAAGAGGTGAAGAAAAAGACTTTAGAGTCAAAG

SEQ ID NO: 973 ACGGGGAACAGGACAAAGAAAAGGAACGAGAAAAAGACAGATCCAAAGAG  
ATAGATGAAAAAAGAAAGGATAAAAAATCCAGAACACCACCCANGAGTTACAATGCATCGC  
GAAGATCTCGTAGTTCCAGCAGGGAAAGGCGTATTGAGGAGGAGCAGGAGTTCTTCCAGATCGCC  
AAGANCATNNNAACCATAANANGGAAATTTTCTATATCTCCGTCCCCCAGNAGCAGAAATAANA  
ANGATTAAGAGAGANAAAGAANGGGACCACATNATGAAAA

SEQ ID NO: 974 GGTACTTTTCTTCTGTGCCTACTCAGTAAAATTGAGAAAATAATTTTGTCCCA  
CTTCTATAAAAAATATCACCCCATTTGCTACCTAGCATATGAGAGAGACAAGACTAGCCTGTCACGT  
TTGTCCTTAAAAAGAAAAATGATAGGCCTGGTGGAGAGCAAGCTATATGGAAACATCAGCCTTTC  
CTTTTCTAAGATACGTCACTCCAATTCTTCTTAAATATAAACTTCTCCTTGATTTATATTAACCC  
CTTAGGAGGCACTATAATAACTTTATGTTCCCATGGAACATAGGGTTTTCAAGAATTTCTTAAAA  
TAAAAAAGGTTTGAAAGCTATGGAGTAAACTCAGCAGCTGTGATATTCCGTTCATCTACCCACAG  
ATGTCAAACGGTCACTTTTACAGAAGTCAAAGGAAAAACAAACCACAGTTCTGCACTCTTCCAC  
AGGGCAAGGAATTGATCATATTCTATTTTCTTTTTCAGGAATTCATAAAATCCC

SEQ ID NO: 975 GGTACGCGGGGGTGTCCCGTTTCATCCAAGGCGCAAGATGGCGCTGCTTTTTG  
CACGTTCTTTGCGCTTGTGCGCGTGGGGAGCCAAACGATTGGGAGTGGCCTCCACAGAGGCCAG  
AGAGGCGTNAGTTTCAAACCTGGAAGAAAAAACCGCCACANCAGCCTGGCANTGTTAAGAANA  
NNATACCNNGGTCAAAATATGGCTTTGGTGGGATTTGGANCCCAAGGNGNCTTGAATGTGGA  
GCGCTTCNNGGAGTGGGCATTGNTGCTGGCANACNCAGCNGTCAACNNNGGNGGACACTACTNGG  
AAGTGACAGTGAAGCGCTCCCANCAANTTCCGGATTGAGTGGCATANGTGNACATGTTCCNGGAT  
TNNTTCATTGNTGTTGANGANCAGATCTTGGGGTGTTCACCTTATTNCCNTCTCNAAGANGGGAC  
AAATGNNTATTGCGATTANTTNAAGGGTTAAAAACNTTNTNAAAAGCAATTTGTNGTGAACAA  
NGTTTNTT

SEQ ID NO: 976 GGTACAGAAGATAAACTTTGCTATTTTCTTTGGCGAATCATCCAGGATACCCCT  
TGGACATAAAGTAGTTCACTCCCTCATCTGGGTGGCATTAAGGTGAGGCTGCCTTCATCCAGCT  
GCATATAACAATTTCTAAAAGAAAAATCTAAAGGTGGGTCTTATTGGATATGGGAACAGTGACCC  
CAAGTGAGATTGCACAACCCTTGCCAGAGAAGTTTCATCATTCGCAAGGTCTCGCCAAACACATGA  
AGCCAAGCAAAGGTCAAGTTCAGTTCAGGTAGGACAAGATGGTAAAGCTTAGCTCAGGAGGCAAC  
ATTTCCAAATTAATGAATCCTTCTGTTCTTTTCGATTTCCTTGCCTTCAAAGATGATATATGTCAA  
TGCTCTCCTTGACTTGTTTACGATGATTGGTGTAGAAATGTTGCTCGCAGCCATTCTCCTGCTCTG  
CTCTCTGGGGAAGTAGCCTTGCTCACTGTAGCCTTCTTGTTC

SEQ ID NO: 977 ACAGATACGCTGTCCCATACATCAGGATCAAATTATTAGTTTTCAGTTTCACAT  
TGTAGGAATTTAAGATTTTTTTTTTTTTTAAACACAGGAAATAATCTCATCATTTCCAAAGATGTCTT  
CATGTCCCATCAATGACATGCTACCAGACATATCAGATCCACAGGATAATGGGCNCCAAGCTAC  
CCAAGTAGATGTTTCTGGTATTCTAGACTGCCGTTTCATGCTTGTTCCTAAAGTATACTTAAAAGTT  
TCAAATACAGTTTCACTTANAACTGCAACCCTCCAAGTAATGTTATGTTTACTTAGGTATTAATG  
TTATGTTTACTTAGGTATGTATCAGAGGCAATAATTTCCAAAGCAGATCTTAGAATATAACCAATT  
TGTTAGATAACTAATTTATCTCTATCACATCTGTTTACAAGCAAAGTATTACTTTGNCTGGACTTA  
TTTCATCTTCANGGGCTGGGAATCGGGGGCAACAAAACCAAG

SEQ ID NO: 978 GGTACTCGCCCCCTTTTGGAAAGACGACGCCCCAAGGCACCCGTTGTGTGCACAC  
GTATCCGTGTGTGGACCTCTGTCCACACGTGTGTGTTTTGAAAAGCCTTTCTCACCACATGTTAGG  
ATCACAAGTGTTCAAAATCTAGGCACACTTTTCAGAAGTTATTTCTGGAAAAAATGTTGAATTCTC  
ATAACAGCAATAAAAGGAAAAATACCGCCTCAGCCAATAATANAAAAAANNNAAAAAAAAAANA  
AAAGGTCTATGTGACTATCATTGATGCCCCANGACACAGAGACTTTATCAAAAACATGATTACA  
GGGACATCTCAGGCTGACTGTGCCTGTCTGATTGTTGCTGCTGGTGTGTTGGNGAATTTGAAACTG  
GTATTCTTCAANAATGGGCAGACCCGANAGCATGCCCTTTTGGTTTACACACCTGGGTNGTGAAA  
CAACTAATTGTCNGTNNTTAAACAAANTNGATTTCCANCTGAGCCACCCTACAGCC

SEQ ID NO: 979 GGTACATCTTTGGCTTGTGAAAACCAAACATCTTTTCTTCTGGGCAATAGTAG  
GCCTGGTCTGAAGTCTTAGGTCAATTAGCTGTCCACTCAGATGTTGTCAACTTCTCGTCTGGTTTC  
AGGGTTGAAGCTGTTTGGTGAATCAGAGACAGGGATCTTTGGTGGGAGTCATGGGCCATTTTCAGG  
AAGTTGTGCTTCCAAAAGCAGCCCGTTCCACGCGCAATAATGTCCCGAGGTGCGGAGTGCACG  
CCAGGCCAGTCCCCCTCAAGCGCTCCCTCTCAACTGCCTTGTCTCCCGCGT

SEQ ID NO: 980 ACTTTNTTTTTTTTTTTTTTTTTTTTTTTTAGATTANAATAAAAAATTTATTTTT  
GTAAANAATTATATTTTGTATTTGCAAAAGCTGAAAATGCTCATAAAAATTACCANCCCANAGCTT

GGATTTCCACCGGATCCACCACGTGANACAAAAGAGTNTGTCACTTTTTTTNTTNCAGGGTTGA  
 GGGCCTTTTCTANACCTTGATGTGTTTTGAGGGAGCTGATACTTTCAAGCAATANCCAGCCGA  
 GNGNGGACCTGGTTTCCCTGGATCTGCACCTGAAGGCTGTCCTTGGCCCCAGGGGCAGGATTG  
 ACGTGGTGCTAGCCTGGCATNGCTGCTGAAGGATGGCAGCCACTGAGTATGGGGTNCAAACCAT  
 NGCCTTCAAAGTTNCGGACCACGGTACCTTTTTATTAAACNCTCTTTTTGCTAGGGNGGATGTCA  
 ATTGGACANATTNTTCTTTNTTANAATGGGGCTNTTTNTCCCGGGAG

SEQ ID NO: 981 ACGCGGGGAGAAATTAGACACCTTCCCACTGGGGACACAGACACAGGAAGT  
 AGCAAGAAGGGAGATGCCAAGTGACAATCACCAGGAAGATGCCTCTCTAGTGACCTGGGTAGT  
 TTGCACGGTTTGGCTGGAAACCAAGTCCCCCATCTCTGCCAGAACCCCCCATGTGGGCCACTG  
 TCCTCAGACAGCTCCTGGAGCTTGTGGATAAGCACTGGAATGGCTCCGGTCCCTCCTCAACA  
 AGAAGTTTCTCGAAAGTTTGAAGCAAAAACCTGGTCAGAGTGTGGAGGAAAACCTCATTTTGTG  
 AGAAAAAATTCAACAGTTGGAGGAAGGTGCTTGCCATCTCAATTGTGAGTGGGGCAACAGTCACA  
 TACTTATGATGATCTTCTGCACAAAAACCAACAGCTGACCATGCAGGTGGCTTGCCTGAACCAGG  
 AGCTTGGCCACTTGAAAAAGCTGGAGAANACAGTTGCCATTCTCCATGAAGTCAGAAAT

SEQ ID NO: 982 GGTACTTTTTTTTTTTTTTTTTTTTTTGAATGTCCCAGGCTGCCTTCCTGTCC  
 TCCCATCCTGCATCCTGCTTCTGTTCCCTGGATAACCGTAGGATGGTTTTATTTCAAGTTCATGCACA  
 AATTAGTCTGGACACTGTGGAGTCATAACAAGAGTGGGATGGAGGTTCCAAGGCCAATCATTTTC  
 TTTGGAGGGAAGCTTGAGAAGTAGGTAGGCAATGCCACAAAAGCCACTGCTTCTCCTTCCCAT  
 CCCCAGCAATGAGCTAAGAGCCAACCTCAGCCAGAGAGAGGGTCTTTGATCACAGAGGGTCAGTC  
 AGTGGGATGTGATTAGTGGGTTTGGGGATGGAGTCCCAGGTAAACCAGAGGGCAAAGGGATCCTC  
 TGAGTCTGGCTTAAGGGAATGGGGTTTCAAGTTGANGGTTGGGGACTTAAGGGCAGGGGGANCA  
 NCAAGGTGAGGAAAACANAAAGGCTGGCATGGCGAAGCATCCTTGGC

SEQ ID NO: 983 ACTTCCTCTAACATAGGTGGTCAGCATACTCTTTCCTCCTTATCCAATTTTAT  
 CAAAAATAGAGATTAGGGCAATATGACATAACAGCATTAAAGCACTAAACCATTTGACAGGGATT  
 CTTTCAAAAGTGAAAACCAAGTGTACC

SEQ ID NO: 984 ACTTTTTTTTTTTTTTTTTTTTTTGCATATTAAAAAATTGTGCATTCCAATA  
 ATTAAAAATCATTTGAACAAAAAATGGCACTCTGATTAAACTGCATTACAGCCTGCAGGACA  
 CCTTGGGCCAGCTTGGTTTTACTCTANATTTCACTGTGCTCCCACTTTTCCCCCACTTC  
 TTNCTTACCAACATGCAAGTTCTTCTTCCCTGCCAGCCAGATAGATAGACAGATGGGAAAGGC  
 AGGCGCGGCTTCGTTGTGAGTAGTTCTTTGATGTGAAAGGGGACAGCACAGTCATTTAACTTGAT  
 CCAACCTCTTGCATCTTACAAAGTTAAACAGCTAAAGAAGTAAAAATAAGAAGGCAATGCTTGT  
 GGAATGTCCCCGGGGAAAAAGGATTCTGGATGAAATTGAAGAACATACATCAAAATCTATCCTTA  
 CCTGATGCCGAATCAGATGAAGATGAAGATTTTAAAGACCG

SEQ ID NO: 985 GCGTGCGCGGCGGAGGTACGCGGGGAGGAAGGAAATTGACGAACACGTGA  
 CGCGGTGCGGGCGGACCACTGCAGACTGAGCGGTGGACCGAATTGGGACCGCTGGCTTATAAGCGA  
 TCATGTTTCTCCAGTATTACCTCAACGAGCAGGGAGATCGAGTCTATACGCTGAANAAATTTGACC  
 CGATGGGACAACAGACCTGCTCAGCCCATCTGCTCGGTTCTCCCAGATGACAAATACTCTCGAC  
 ACCGAATCACCATCAAGAAACGCTTCAAGGTGCTCATGACCCANCAACCGCNCCTGTCTCTGA  
 GGGTCCCTTAAACTGATGTCTTTCTGCCACCTGTTACCCCTCGGAGACTCCGTAACCAAACCTTT  
 GGACTGTGAGCCCTGATGCCTTTTACCAGCCATACTCTTTGGGCATNCAAGTCTCTCTGGCGAAT  
 GATTATCTTGTGTGAAGCAATCATGGTGGCATCCCCATAAAGGGACACATTTGACTTTTTTTTAT  
 ATTTTAAATTNCTTNCGAATNTTAAAGATAAATGTTTCTCNGNANNANCCTNCCNTTNCNNAAC  
 NTTCTCCCCGGGGGCCTTTAA

SEQ ID NO: 986 GTTCCTGAATGCTAAACTGCCTGGCTCCCAGCTTTTTTCATTAAACTTTTCAGG  
 GTCTTGGTTTCTTTATCTGTAAATGACAGAGTTGGACCAAGTTAACTTTAATGGCCATCCTTTTACA  
 CCACACAAGTTGATAAAATTTATCTGTTTCAGCAAAGAGATTGAACAAAAAAGCACGTTAGTAATA  
 TGAAGACAGGAAAACGAATGAAAGTCTAACACATAACTCATATTGATTACTTTATTCTGTTAGA  
 TTTTACACTCTGAAAATTTACCTCATTTAGTTTGTACATTATAGCAAAGTGGTATTTATTCCAGGA  
 ATGCAAGGTTAGTTCAACATCTGACAATCAACCAATAGATTACCAACATAATAAAAAACAATAAAA  
 GTAGAATAGAATAAATATTAATAGAATAAAAAAATCACAGGATCATCTCAATAGATGCAGAAAA  
 AATATTTGACAAAATCAAAATACCCCTGCTTCTAAAAAGAAAAATATTTTNGGAATAGNAATGAAC  
 ATACCTTGATAAAGGGCTTTATTTAAAAAATCCAGCTAACCGTATACTTATTGGTAAAGACTGG  
 AAAGCTTTTINCAAGTT

SEQ ID NO: 987 TGTACGCGGGGACAGACACTTGTAAGGAGGAGAGAAGTCAGCCTGGCAGAG  
 AGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGCTTCAGCCTGAAGACAAGGGAG

CAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCTCTTGGCCTCCAACTTGTGGGCT  
 ACATCCTAGGTCTTCTGGGGCTTTTGGGCACACTGGTTGCCATGCTGCTCCCCAGCTGGAAAAACAA  
 GTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTG  
 CCACACACAGCACAGGCATCACCCAGTATGACATCTATAGCACCCCTTGGGCTGCCCCGCTAACAT  
 CCAGGCTGCCAGGCCATGAATGGTGACATCCAGTGCAATCTCTCCCTGGNCTGCATTATCTCTG  
 TGGTGGGCATGAAGAATGCACAGTCTTTTGCCAGGAATNCCGAACCAAANACAGAATGGCGGTAG  
 CANGGTGGAGTCTTTTATTCCTTGGAGGCCTCTGGGATTAATCCTGTGCNGGAATNTTATGGGG  
 ANCTCNNGGACTTTA

SEQ ID NO: 988 GGTACGTTTTTTCCCATATAGGACATACTTCGAAGTAACAAAAATAATTTTT  
 TAATCGACTGTAGTGTCTTCTTCCCGTTAGAAGTAGTTCTTTTTCCCTCATGGTACCAGACT  
 GGCAATGGTTCCCAAATTATTTATGAGATCAGCTTTGGTCATGCCAATGCCTGTGTCTACCAAAGT  
 CAGGGTACGCGGGGATTCTTTTCATAAAAGATAAATTATGGACTAAATCAGGATCTAGAATCAGAGT  
 CAGTTAATCCTATTTTATCCCTAATCAATTTTTAAAGATAACATGGCATATATGTGCACATCTCA  
 GCAAACATGTAAAGTACCAAATACAATCTGTAAATAATGGGCCATTTCTTCCCAAGCCCACAGCC  
 AAGATGAAGTGTAGGCTAAATTCAGAGCCCTGGCTCTTNTCTCAGATGAAGTGGAAAACCTGCGC  
 ATGACANGGCCGGTTCCCTTCTCAAAGGNTTCTTCCCTCATCCATNCCTAACAAAATCGCCGAG  
 GGGTCACGGCCCNANAAAGCNCAACATGGGTTCTGAANTCCAAAAAAGCCGTTGGTGGGCCCAA  
 NAAAAANA

SEQ ID NO: 989 GGTACTGGCTTGAACAAAATTTGTTTTGTGTGTTAGAGTTATAAATCATTAA  
 TTTTATTTTCGGGTGGTTTACGTTTATGCCAGTTCCTTTATATTTAAATTTCTTGTTTTATATATTTTG  
 AATGTCTTTATAGATTTCTTTAAATTTCTTATAGAACCATTAAATAGAAAATCATTACATTTAAAT  
 ATACCTTACAGCAAAAAGCATCCAAATAAGTATAGGGTTTATGTCTTATTTTTCTTTCAGCTGAAT  
 ACGAATGAACACAGTGGTGGAATTTCTGAAGGGAAGTGATGAAATTATATTTATTTTCAGTGGGCA  
 CTTTTCCATTTTACCACTGTACAAGATGTCCAAATATTGCGAAGATCTATTTGGGGATCTCCTGTTG  
 AAACAAGCACTTGAATCACATCCACTTGAACCAGGCAGGGCTTTGCCATCCCCCAATGACCTCAA  
 AAAGAAAAATACTCATAAAAAACAAGCNGCTGAAACCTGAAGTTGAAAAAAAACAGCTGGAAGC  
 TTTGAGAATGATGGAAGTGGAGAATTGCCTCCACAACATTTAAAGGCGATATGAAAAGAGAT  
 CGAAG

SEQ ID NO: 990 GGTACAAAGTTGTCCAGTCTTAGGTGCTGAGGCCAGTATGGGTGCAAGGGGC  
 CCGCGTTGGTAGTCATGTCTTTGTGGGCTGATGGCTGCGTGTGTATAGGCAGGAAGTTAAAAAA  
 AAAAAAAAAAAGGAACAGAAAAACCAAGGCGCGCGACGGCAGCGGCCTCCCACGCTGGCTCCGG  
 GCTCTCTTTAGGCTGGGTCTCGCCCCCGCTACTTTANTTTTTTTTTTTTTTTTTTTTCCATATTTTCT  
 TCAGCAGCTTTTCATNTGGANAAGCCGTTTCCAGGATATTTNTTNTGATGAGCTNTTTTCCCATATA  
 TACTNTTCCATGTGCGCTTCANTCTTACTTTTTTCCCATNCTTCAGTNTTCCAANACTTATGGGATA  
 ACCTGTAGTAGATACACTGCATAATCGNAATTTTGCTCCTTAAAGGGGCACCTTTTTTTTTTTTG  
 ATGGNTAACCCCATTTGGGGNCTTGCAGTATCGCTCTCCCANCAAGTTAAANAAATTTTTTTTTTN  
 GNNGGNACATTTTAAATTAAACNNGGNAAGCTTTTCTNTTA

SEQ ID NO: 991 GGTACTCACAGCTGTGGAGTAGGGGTATGAGCTCCAGTATTTTCAGCCAG  
 CAGCCACATCCATAAAGAGCAGCCTGCCAACTCTCCCCGATGTTTCAAGGCCAAGCCTCCACT  
 GGAGACAGCAGCAGCAACATTCCCTTCGTGGTCCACAACCACAGCGCCTACCGTGTCCAAAGTGC  
 CTGAGTCATTTTCTTCTCACTTGATTGNCTTCTTTTCTTATAGTTGCATAAAATCTGTGTCCACCCTT  
 TCTGCCAGCTCTAGTTTCTTCTTGTCTTTTAAATGCAGCTAAACTGAATCTTGTGGTCATGATGT  
 TAGGAGGGCAAAAGGGGTATTCCATGATCTACTGCCATCTGTAGGCTCCTTCTCCAATAAAAAG  
 CAGGGGAGGAATTTTGCCAGCCGAGAGCTTGCCCTTTTGCCCTTCACATAAAAAGTCTGTTGGGC  
 AACCCGANACTGGGTTTTTGATTTCCCTCAGTGGCTCCAAGTGTCCNAAATTTAANGGATTTTC  
 CTTCTTATGCTNGNGTTACACCTCAATTTTACCTTACAAATTTAAA

SEQ ID NO: 992 ACAACTCTTGCTAATGGAATGCTATAATGCACAAGGTCAAGGATTTAATAAA  
 TTCTAAAAGTGTCTACATATATCAGTGATAACTGTATTATTAGAAATATAAATGTATAGAAATATA  
 AAGTATATGGTATTAAAAACAGACCTTGCTAATATAAACATATATAAAGTATGTCACTTCTCCTGT  
 AATAACAGCATAAAGATCGATCTACAGTTTGCCCTTCGCTGGCACTCTTAAACCACTCCTCCAAT  
 GGTCAATGTTGACCTTGAATCAACAGCCGCTGAACCCAGGAGACCCACAGATGTGTAGATTCAG  
 CACCTAGAGGGCCCCCTACCTCTGTGCTGTGTGTTCCCATGACTCCAGAAATAATTAATCGCAA  
 CTTGCATTATTAAAGTCCCAGGCAAGTTTGAAATCTAACTAGAAAAAAGTANCAGCCAANGCAA  
 ATACCGCNGGAATTTGGTAGAAAAACCAAGAATTCTTAAATGCTTTCCAGTCAAGGCAAAA  
 TTAAGGGGAANTNAGNNCCNCCCCCTTCC

SEQ ID NO: 993 GGTACTTTTTTTTTTTTTTTTTTTTNGAGGTTTCAACTTAACATTTATTGCACAA  
CTCACTTCAGGCCACTGCTGACATCCCAAAACACAGCCTAGTAACACACAACCTTCACCTTCAAGG  
ACTGCAGAGATAGACCAACTATATATAAATAATCTCGTGGAACGTAATTTGGGTGAGAAAAATA  
TGACGCTAGAGACTTTGCGATGAATATTTTGCTAAGTTCAGAAGCTTGAAAATACAGTTGGCCTA  
CAATCTNAAATTTACCCAATAACCGCGGCCCTTAGCACCACACCTTCCATCATCTTTTTACCTCTG  
CTGTTCTTGCCCTTTACAGNTGCAGCTCTCTCAACTCCCTTGCCCTAGTTGACTCCTCTTTCCATTT  
CCTACAACCTTTGNGGCGGGAGCATTTTAAANAAAAGCCATTTCAAAGGCGTTNAAAGGCCAACTGN  
NCCAAGGCATTTACTCAAAAAATTTATTCCAACAGGGATCTTCAATTTTTTAAATTCTTGACCTTTT  
TTTTGCATAAAATTTAAACCAAGCTTAANATNNGGNAGTNGNAAATTTAACTNCGTCCCTTAAAT  
TAAACCTTAAATTGNTTTTNAAAAANTTTATTACAGGGAGTACNTTGNTTTTTANATGGANTTTGC  
CTGNGGNA

SEQ ID NO: 994 ACNAGCATGTTTTTATAGAACAATGTGCTCACTTTGAGAAATGAGAAACATG  
AATTGCAAAAGAATATCTCATATCCTCCATGTTTCAATCTGAAGCAGCACACAAGGCTACTTATTT  
TCTCCATTAATGCCAGATAGAAATTAATAACCTTCTGAACCAAGTGCGGAAATCCAAAACCTG  
TTGAGATGTTTGTCAATTTATGATAGAACAAGGTAACCTCTGGTAGGCATTTGGAATTTNACAGTAN  
GAGAGTGATGACTATACGTAAATGGTAAGTCAATAAATTATGNTGGCTGAGTATCACTTATCTAA  
AACTGCTTATAACCTGGTAGAAAAGCTGTCNATTTATGTAAANATTTAAAGAAATGTGTTNAAAA  
CAGGAATGCAACACCTGAAAGTCTGNTTGCTNACTACCATTTTNAAGGGCCTTCCCTTTCCAAGC  
TCACCCCTAATGCTCAANAAGANGCNCATTTTAAAAACATGCTGGGTGATGTAAAAAACACNCGA  
AAAAAACCTTTCCCTGGTTAAAAATTGCATTAAGNTTCTTNTTGACCAAAATTTGNNNAGCTTCAA  
ACCACTTAAGAAATAAATGCCTCTNTTTCCAAANAAAAAANCCANTNCTTTTTTNGGGGANCNCAG  
NCCTTTTAACTTANNCGANGNANCTTTTTT

SEQ ID NO: 995 ACGCGGGGGCCAGGGACTCGGGTGCCTGGGGCAGACGAGGCCGGCTTCTCC  
GCGGACAGCTAGGGGAGAGTGTCTGGGTGTCAGCCAGAACATGTCTTTCAACCTGCAATCATCAA  
AGAAACTGTTCAATTTCTTAGGAAAATCACTGTTTAGTCTTCTGGAGGCTACGATTTTGCCTTACT  
CCCAAAGCCACGGAAGAACGTTGCTGGTGAAATAGTCCTCATCACAGGTGCTGGAAGTGGAAGTCTG  
GAAGGCTCTTAGCCTTGCAAGTTGCCCAGGCTGGGATCTGTTCTTGTCTCTGGGATATCAATAAGG  
AGGGGAATGAGGAAACATGTAAGATGGCTTCGGGAAGCTGGAGCCACAAAAAGTGCACGCCTAT  
CCTGCGATTGCACCAAAAAGGAAGGAGTGTATAGAAGTAGCCGACCAGGTTAAAAAAGAAATC  
CGGCGATGTTTNCATTCTAATCAACAATTGCCGGAATCGTAACANGCAAAAAGTNCCTTTGACTGNC  
CANATGAACCTTATTGNAAAAGTCTTTGATGTGAATTTCAAAANCCAATTTATGGACNTATAANNCTT  
TCTANCTGNTATTATTGCTAATGACCATGGCCTTTTGGTTGCTTTTCAAGTCAACTGGATAAATGGA  
TAAATGGGTNGCAAATATTGGCANGNAATTGC

SEQ ID NO: 996 ACGCGGGGTTGGCAACGAGGGACTCGGCCTCGGAGGCGACCCAGACCACAC  
AGACACTGGGTCAAGGAGTAAGCAGAGGATAAAACAACCTGGAAGGAGAGCAAGCACAAAGTCATC  
ATGGCTTCAGCGTCTGCTCGTGGAACCAAGATAAAGATGCCCATTTTCCACCACCAAGCAAGCA  
GAGCCTGTTGTTTTGTCCAAAATCAAACTGCACATCCACAGAGCAGAGATCTCAAAGATTATGC  
GAGAATGTCAGGATGAAAGTTTNTGGAAGAGAGCTCTGCCTTTTTCTCTTGTAAGCATGCTTGTC  
CCCAGGGANTAGTCTACCAAGGTTATTTGGCAGCTAATTCTAGATTTNGATCATTGCCCAAAGTGN  
ACTTGCTGGTCTCTTNGGAATTTGGCCTTGGAAGGTATCATACATANGAAGTATGCCAGAAGTAA  
ATCCATTTTTTTGAANATCACTTNCNGGGGGCTTGTTTTGGTCCACAGCATAACAAGGACTTGCTT  
CTTANCTGGGANGGATGCAAAATTAAGCATGGATTAGTGAGAANGGANCTTTNACCTTTAANTT  
CTTAAATCGGGTTTGGATTTCAAAGTTTTAACTTGAATTNATCCTGGCCGGACNCNTTANGCGAA  
TNNACNTTGGCGGNGTANTANGNT

SEQ ID NO: 997 ACTAGGTGCTGCAATGCAAAGGGTTATGACAAAACCTGTCTGTAAATGTAGGA  
TCTGAATTGGTCTTTGATAGTTTTCTGATTTGAGAAAGAAGTCTCTATTTGGCTAATTTAGGGAGC  
TAAGAGAATGGACAAGGTCTGCTTATGTCTTCTGGCAGCCTAGCATAGCTTTTGCTCCCTCAAA  
TCAGTTATAAGTCAAAACAAATAAGGCACATTTTTTAAAAAAATTTCCCTTTAATTGACCAAG  
TAAAGCCATGACATTTTCAATTGGTAACCTGTTTAGAATTATAAAAAATCATTTTATTTGGCCACCCAT  
ACTGCCCAAGACAAAACTTNCAGACAATTCTGATGCCATCCAAGTTTGNCTTACAAACTGCATAT  
TAAAAAATTTTTTAAATCTTACCNNCTAAATGTGATGTGCTCAAGNGCGAACCTATTAAAN  
CTGGACGTGNATTTATTGTNCNTGACAAGCCCTCGANTATTCAAGGCCTTTACCTTGGATATTNN  
GCCANNCCATGGTCTTNGGATATGCTTTCTCCAGGAAATTAACCTCAANGCNTTATTTCCAATT  
TAACANTTTNGGCACCGGCNAANCTGGAGGACNCTTGGNTTTTGGTAAACNCCCGGANCTTGG  
CGNGAACNCTTANGNGAATTTAAACACTGGGGGG

SEQ ID NO: 998 GCGTGGNCGCGGCGAGGTACTTCNCAAGCAAGCCCCTATGATTTGTCACTAT  
AGATGGAACCCTGACTTCTGCCCATCCCTTCTGCCCAACCTAGAACCCAGGCCTCAAGTCTTTA

CCCCACCCCTTTCTTGTTCTTCCAAGAAGCAGATGCCCAGTTGCTCAGGCAGCAGCGGTAGAGACT  
TGAATCTGCCACCAGTCACAAGGCGGGTCACAGATTCTCTCTCTCTCTCTCTCGTTCCTCTGA  
ACCCTCCACCAATGTGCTCAGCCTGTGTGCTGTGTGGCAACAGCATTCTGGTTCCTACTGCCAAG  
ATCTCCCACCACTCTGCTGGGATCTGCAGNGGCAANGGAGTGGGGGTGTGTAAAGGGGAAGTCA  
TCTTTTGAGATCCAGATAGACATGGGTTGNGNACTTTACGTCCAGATGGGGAAGCATTCTNTTCTG  
NAACCTTAAATAATCATGCANCTNTTCAGACNGACGCCATNGGTCGAAGGGCTTAAAGNGGA  
GGAAGCAAANNNGGCAAGCCCTGGCTTNNCGNGGANGCTCTAACTTTTNGATTCTTACGGTGCAA  
AGGAATTGGGGTTTTNTTCANGCCTNNGGGGCCAATTGAANTGGGTCTCTCTTAANTGGGCCATGCC  
AAACGTTCCAANCTGAGGAAATTACAGGGGCCNCCTGGTCTGANGGGGNNCTAC

SEQ ID NO: 999  
ACNTGGCTTGCAGTTTCTTCAAAAAACATTGACATCACAGAACACATAGATTT  
TGCCACCCCTATACAGCAGCCAGCAATGGAGCCTCTTTGCAATGGCAATCTCCCCACGAGTATGCA  
TACCCTGGACCACTTGCATGGGGTTTCCAACCGAGCCAGCCTGCACTACACAGGGGAGAGTCAGT  
TAACAGAGGTATTACAAAATCTCGGCAAAGACCAATATCCACAACAGTCGCTTGAACAGATTGGC  
ACCCGAATTGCCAAAGTTTTGGAAGAAACCAGACGTCTCTGGGTCTCTCCAGCATGGCAGCCCT  
NTANTGGAGGGTGAAAGGCCAAGGAAAGAAGCAATCGACTGCCTCCGCCAGGCTCTGCACTATGC  
GCNACACCAGATGAAGGATGTGCCCCTGATTAGCCTGGCCACATNTTGGACAATGCNCAACTNTN  
GAATGACGCGNCTTAGTAACCACTATGGCAGTAGANANTCGCACCACTTTGNTGGGAACCAC  
TTTACTCTGGGCAATNGCTACGTGGCAATTGGAAGAAATTGNAAAAANNCTNGGTGGGATGAAATC  
ACNTTGAACCTTATGCCGNGTTGCCACCAAACCAGATNCAACCTTCAGGNACTTAANNNGAAGAA  
GGACGGGCTTTNTAAGGACTTTTCTTTTTTTT

SEQ ID NO: 1000     ACTATTTAATATATTTCTCCATGAACTTTTGTGAAATTCAGATCGCAGTGTGT  
CATTTACAAATCTTTTGTCTTTCTTCTGGTCATCTACACCTTTTGCACAGTTCTTGAAGACAACGTC  
ATCATCCCACCTTCTTTAACTTTGAAGTTGGCCTGAGGCTGGGATGGGCCAGTGAGATTAAAGGAG  
AGGGTTTCCGCTCAGAATGTTTTCCATACGAATCCTCTCTTCTTCAGCTTTTTGTCTTGTTCCTTCC  
TGGCCTGCTCTTCAGCTCTTCTTTTTTAATTTTTTCCAGTTCTGCAAGAAGAGCTGCAGTATCATC  
ATCATCACTTTCTTCTTCAAAATCTTCATCTTCTCATCTGGTAAGAGGGGTCATCTGCATTAAAGG  
TTGGCGGCANGAATCTGGNCTTACCCGTGGCTTTTTTGACACTTGAAGAAGAAGGTGNATGGTCTC  
NGGGTNGACNANTCCTATTTTTTTCTTTGGCAGAACTNTTTTTTTNTTCCACTNTTTTTCTGAAGC  
ACGGTACCAACCTTTTNAGGGCCTTCCGANGAAGCCGGCTGGATTATCTTTGGTGAAAAGGNAG  
CGTTTTCTTGANTCTGNTTTGAAAGTGTTAAAAACNTTCCTTTCCCTTCNCTTGGCGGGCNAAG  
GTGGCCNGTTNNNTGT

SEQ ID NO: 1001      ACGCGGGGGATCAAGTTTAAATGACTGTGCTGCCCTTTACATCAAAGAAC  
TACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGCAGGGAAGGAAA  
GAACTTGCATGTTGGTGAAGGAAGAAGTGGGGTGAAGAAGTGGGGTGGGACGACAGTGAAATC  
TAGAGTAAACCAAGCTGGCCCAAGGTGTCCTGCAGGCTGTAATGCAGTTTAATCAGAGTGCCAT  
TTTTTTTTTTTGTTCAAATGATTTTAATTATTGGAATGCACAATTTTTTAATATGCAAATAAAAAAG  
TTAAAAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGTCCTNGGNCGCGACCCCCCTTAAGG  
CGAATTCAACACACTGGCNGGCCGTAATAANGGATCCNAAC TNGGACCCAACTTGGNGTAATNAT  
NGGCATANNTGGTTTTCTGNGNGNAAATGNTATCCCTTACNATTTNCACAAACATACCAACCGGAA  
CCNTAANGGTNAAACCTTGGGGGCTAATGAAGNGACTTACTCNATTAATTGGGTGGCCTACTTGC  
CCTTTCAAAGCNGAAGCAACTGCNGGCCACTTGATTAAGAATCGCCACCCCNNGGNAGNGGGTGGGN  
TNGGGCTTTTCNTTCTT

[illegible]

SEQ ID NO: 1003      ACTGACACATCCAAAGCATGAGTGTGTGCAGAAATCCCTTGTCTATTCTGTCT  
GTATAAAGTGTTTCATTATGACCAGATCTCTGATTGTATGGTCACTAGGTATGCAATCACGCATT  
AAAGAGGCTCTCTTACACCATCACTGTGATTGCTCTGAGAGTTGAGGGACATATTGGGCTTTATTTGG  
ACAAACCAAACTTTTAGCCTGAAACCACTTTATGCCACTAAGTCATAGCCTAGTTGTGCCAGTT  
ATTTGTCTCTCTGAAAATGCCTGAAACATCAGACAGACAGTCTTGCTTTACCCAAACTGATCAAA

ATCTTTAGGAGCACAAATGAATTTTTTAGTCTGAAATACCAAATAATGAATTGGTATACCATATCC  
GGAATCACACATGTTATCTTAAACCCACCATCATACCTAAGCTTTTGGCCAAACCTCTTATAAGTA  
TATCTAGCTGAACCTATTTGGCATTTCATGGGGAACAGGTTCTAAACCTAAAAAGGGGNCANGN  
TTGTTTTAAANAAATTTNTTTTCTTTANGGCCCTNGGACNTTTAATACNNAATTCNTGGAACN

SEQ ID NO: 1004 ACGCGGGGGTGCGCCGGTGGCGGGACTCTGGGGAAAATGGCTGCGTCTTCGA  
GTGGTGAGAAGGAGAAGGAGCGGCTGGGAGGCGGTTTGGGAGTGGCGGGTGGTAACAGCACACG  
AGAGCGGCTGCTGTCTGCGCTTGAGGACTTGAGAGTCCCGTCTAGGGAACCTTATAGAAATGCTGG  
CAATTTCAAGAAACCAAAAGTTGTTACAGGCTGGAGAGGAAAACCAGGTCCTGGAGTTGTTAATT  
CACCGAGATGGGGAATTTCAAGAACTAATGAAATTGGCACTTAATCAGGGAAAAATTCATCATGA  
AATGCAAGTTTTAGAAAAAGAAGTAGAGAAGAGAGACAGTGATATTCAACAGCTCAAAAACAGC  
TAAAGGAAACANAACAAATACTGGCACAGCTGNTTACNAAGCGAAGGAGAACTCAAGTCCAAT  
AGAAAAAGCCAGAAAAGGGGCTNTNTCCTNTGAAAGAAATAATTAANTTTGCNCATAGGATCAAG  
GCCAAGGAAATGCTGGATTGNGCCTCCACTGACCTGGGGTTCCANGGGANCNNCCGGAAACCCTT  
NCCACTGAATTN

SEQ ID NO: 1005 GGTACACAAGCCATCAATTACCACTCTCTGTCTGGCCTCCATGAAAATCCAA  
TTGGTCTTAAAGCCTGTCTAAGCAGAGAGCAAGCAGACCACATCATCGTGTTCAACTTGAAATTC  
ACCCCATCCCAGCACCCACACCCTTCCCCACACACACACACACACACCAAAAAAAAAAAAA  
ACAACAATTTTTTTTTTTTAAAGAAGAAATGGCCATTAAGGTCTGGGCCAGTTGTAAAGGTTGNN  
TTAATCCTGCAGCCAGCACTTTTAAATAGCTACTGGCTTACAGCAGGCCAGTTGGAACNTTGA  
AGCATGACAGGTAGAGAACAGAGGATGTGCTAAGAACCTGGCCTTGTGTGGGCCTCGAGGCCTC  
CATGCATGCTTACACACCAACATAAGTATACTTTCCCCTTCTCCTCCTAAGCAATTGGTCTNCATT  
GCCTTTAAAGCCTGNACACTCAACTTTTTNAATGGTATTTTANAGATGGACAAACCTTAATNTAA  
CTATGCAAGANAGCATTTTTTTNTAAAACTAAGGNCNAAAGNAAANANANGGCCTTTGT

SEQ ID NO: 1006 ACGCGGGGGTGCGCCGGTGGCGGGACTCTGGGGAAAATGGCTGCGTCTTCGA  
GTGGTGAGAAGGAGAAGGAGCGGCTGGGAGGCGGTTTGGGAGTGGCGGGTGGTAACAGCACACG  
AGAGCGGCTGCTGTCTGCGCTTGAGGACTTGAGAGTCCCGTCTAGGGAACCTTATAGAAATGCTGG  
CAATTTCAAGAAACCAAAAGTTGTTACAGGCTGGAGAGGAAAACCAGGTCCTGGAGTTGTTAATT  
CACCGAGATGGGGAATTTCAAGAACTAATGAAATTGGCACTTAATCAGGGAAAAATTCATCATGA  
AATGCAAGTTTTAGAAAAAGAAGTAGAGAAGAGAGACAGTGATATTCAGCAGCTACAAAAACAG  
CTAAAGGAAGCAAGAACAAATCTGGCAACAGCTTGTTTACCAAGCGAAGGAGAACTCAAGTCA  
ATTAGAAAAAGCAAGAAAANGGGCTATTTCTNTGAAAAAATAATTNAAGTTTCCCCTNAGGATC  
AGNGCAAGAATGCTTGATTGTGCTTCACTGACCTTGGGTNCANGGGACNNCCGGANACCCTTTC  
CAACTTANTTAA

SEQ ID NO: 1007 ACATCCGGCGAGTAGCTGGCGGTCCCGGGTGTGCTGGTTAGTGTGCTCTGA  
GGGAGGGTCCGAGCCAGCCNGCTGTTTTGCCGGAGGAGCCCTCAGGCCGTAGTAAGCATTATA  
ATGTCTTTCATCTTTGAGTGGATCTACAATGGCTTCAGCAGTGTGCTCCAGTTCTAGGACTGTATT  
NNNTNTTTTTNTTTTTTTTTTTTTTTTCGTANAGATGGGGTTTACNATGTTGGCTAGGATGGTCTCG  
ATCTCTGGTCANAGTCTTTNTGTAAATATCCTTGGAANAAGCAATTTTAGACTGTACTGTTGCA  
AATGCTTTAAGGAAAAACAAAACAACTGCAAGTCTTCTGAAATGAAAAAACTCACCAGGGCTGT  
TTTTNAAAAACAACCCCAACCAGCACTNNAATTATGATGCCACAGGGGCCCTGANAAACCAA  
GAAAAAGTTNCNAATCNAAAACCTGGGATGCTCTTGACTATGGAAATATTGCNGCCCGANCCCAA  
GTAAANACCAAACAAGCNTAGGNCCCGTANTATTTGGGGGGATTTTGGCAANAAAAAACTTGNG  
NGGGTGNTTNGGGATTCCATTGATCCCCCAAATCTTCCGGGATGGGTAAAAGCCCANGGCCNGA  
AAAGTTANGTCTCCCAAAGGAAAAAAATTTGGGGGAANATTGGG

SEQ ID NO: 1008 GCGTGGGNCGCGGGCCGAGGTACTACTTTTCTNACTTTTTCTGGTTAGCCAGA  
ATGTTCCATTAAAGAAACAATAAAAGTTGTATAGTTCTCTAAGATGAAAGATTAGTATATTCAATGG  
CTATTATATTAACCACTTAGTGAACATACAAAAAACTATCCTTATATTAATTGACTGAAGTTAT  
AACATAAGAAATAAGTTACACTACTTGTGCTTGTCACTTAATACTTACAAGATTACTCAAGAAA  
TCAAAATGGCTTCCCATTGCTTGACGTTTGTGTCCAAATACTTCACATTTCAATCTCATATACCTA  
CAAAGAGGAAAAAATCCAAACATACTTTCTTACCTAAAAATATTAAAGAAGGCTAAAAGGCATT  
AGGAATTTTTTAAACCTTGAAAAACAGNGGTCACCTCCAACATGTATTTAAAGTTGNAATTCATT  
TTCTATCCTGNGGCATAACTTCAATTTAAAGTCATCATAAGTTATATTAAATGGGTCTTGGCTAGCT  
GGTGNCTATTTCCCAAAGAGANTAAAGGGAGAACCAGAGNGGGCCCTCCTGTGAACNATGGTG  
GNAAGGATGAGGTGANGGGATCTCANCCGGGGGNGAACCAAGTAAAATTTNTTTAGGCCNGC  
NCNGGGNTCACNCCGGGNAATCCACCCTTTGGANGCCNGGGGTGACNCTGNGCCGGAGTCANA  
CNC



SEQ ID NO: 1009     ACACTCTCATAGAGATAGAGAAGATCTAAAAAGTTGAGACTACTCAATCCAG  
TTAACACAGCAGGAGCACTAGAGTTTGTTCATTTATTCTCTCTGTAAAACAAGCTGTGCTTTTTT  
CTTCTGCCTTTAAATGCCACCCGTGTATTCAAACCATGGCCACTTGATACTTATGTAGAATCCATC  
GTGGGCTGATGCAAGCCCTTTATTTAGGCTTAGTGTGTGGGCACCAATGTGCGAGCATCGNTGTGA  
CTTGTGCTGTATGATTCTCACTGAAGAATTTCCCTTCAGCCAAGAAGCAGTGAGGTCTGGGAATAT  
TCAAAGTCATGTCTCTGAATATGTGTCCTTGACATGCAAGCTTTGCAAAACCNCATCCCCGCTAAG  
NGCGAGGCATNACCTTCTCCAAGTGGTTAAGTTCTTTTAACCCCCAAGATCATTCTTGGGNGATAA  
TATAAGTTTTCATTCTACTTAAGGGATTGGNTANAAAACCAAGAAAGAACCCATTAAATTTTTAGG  
TTTTGGAAATTTATTTATATGGATACTTAAAAGANGATTTTAAACCGGCGANCCCTTAATTGGCNT  
ACNGGGANGGACTGGATTTCTGGTACACTGGGGGGAANAAAAACCTCTTTAACGGATTTTCNCTTT  
TTCCGGGACTTNCCTTTTAAANTTTTTTTAAAAACGGAATTTTTTG

SEQ ID NO: 1010     CGAGGTACTACTTTAGGAACAGTGTTGTAGATCCATTTAGAAAAAAGGAGAA  
TGATGCAGCAGTTAAATCCAAAGCTGGTTTCGAGGATGTCAAGTTCGGGCATATATCAGGCATTT  
AAACAGGATTGTAACAATTATTCAAAAATGGTGGAGAAGTTTCTTAGGCAGAAAGCAATATCAAC  
TAACTGTGCAGGTAGCATATTATACTATGATGATGAATCTCTACAATGCAATGGCTGTCAGGATTC  
AGAGACGATGGCGAGGCTATAGGGTTCGGAAGT

SEQ ID NO: 1011     ACAAATGTGCATTAAACAATTCAGTGACGTAGCTGTGGATCTCTGGATGGCTAT  
GTAAGCTGTGAGAAAGTCCCCCACTGGCTTTGCACTTGCTGCGCACCAGAGGTGACCATCCAGGC  
AGTATCAACCTGAAGAGGGGGTGAATCCACGAGCTGCTCGATGGGCTTAAACCGCCACTCGCA  
GCCTCCAGCTCTTCTACCAAACCAGCTAGTTTTTCCATCCGAGCAACTTGCTGATCATGTTTCACCT  
GTTTAAGCGTGGATGCCACTTGATAGCTAAAAACAGATTTCGAGAGGAATCTCTCAGAGCCATCTT  
CTAACATCTCTCCAGCTCCTTTGGGGTAGGTGAAAAAGGGCTTTCCGTGGTGGGGCAAGGNTGA  
GACACTGAAACCAGATCCAGTGACAGAACTTTCCTGCCCCCACTGAGGAGGATGATGATGAAN  
CACCNTTTTCTNTNTAACAGAGTCTTACTTCACATGCCTGGCTGNAACCTAAGGNCTTTTCTCTC  
TTTTTTTTGAGACGGGGTTCCTNTTGTGNCCAGCTTGNAANNGGGCGAATTGGTTACAAAACTT  
CGCTTCCGATTNAACGAATTTCTGCTTAACCTTCCAGAGCTGGGATANAAGGTGCCCNCTGC  
CTGGTAATTTNNTTTTANAAANAANGGGTTCTTTNGGANGG

SEQ ID NO: 1012     GGTACTTTTTTTTTTTTTTTTTTTTTTTTGTCAATATTTATTGGGCGCCTATTA  
TGTGCAAGGCACTACACTAGGCGCTGGGGAAGATACAAAGATAAATCTGACAGACTGCCCTCAA  
GAGCTTACAGTCTAGTATAGGAGCATAAGTCTCTGGAGAAGATATTTTAAAGTGTAACTAACCTNC  
CCCATCCCACCCCCACAAAAAAAAGAAAAAACCTACTAGACTTGGGTCTTCCATTTACTTTTAG  
GTTANCAGCTTCACGTAAAAAGCATAAATCTGAAAGTCTTTTAAATGCATACTTTTACTGGNAAA  
CAAAATTNATTTTCATTAGAAAAAATGCTGAATATTTATTGCAATTAAGAAAAATNTCANCCGGG  
CACGGGGGCTCCACCTGNTNTCCCTAACACTTGGGAGGCCCAAGCNGGTGGATCCTNTGANGG  
GCAGGAGTTTGAGACCCCCCTGGCCACATANCAAAAACCCNTNTNTNTTGA AAAACAAAAATTA  
NCCNGGCCNTGGGGGGGGGCCCCCTGGATTCCACTTTTCCGANGGGTGNNGGGANGGAAATCCTTG  
AACCTTGGGGGGGNGGGTGCNAAACCNAANNNCCCCCTTTCTTTNNCCNGNGGTANAANAA  
ANTTCTTTTNA AAAATANAAATTTTAANGGAANAAAAACTCCCCCTTTTTANNTTTTTN

SEQ ID NO: 1013     CGAGGTACTCCTTGACAGTTGATAGATTATATATTCTTCCATCCCTCAAACCTT  
GCATTCCACTATATTTATTTTTTGGCAAAAGATGAGCTGTATTTGTTGAAATCTGAGACACTATGT  
TCAATTGGATGTATCTGTTCAAATTTATTTCCACGTGACGTGGAAGTCCTTCGTTGGATGTCACAA  
CACTACATTTAAGGTTGGTAAGGATGACTTGGAGGTCCATGGTTTTTCATTACCAACATTTTAAGAT  
TCTGAATGTGATGGAGTCTCACTGAAAGAGTCACCAAAGGNGCCTGCCTCCCCCTGCTGGGAA  
AGTGNCAANTGGAGACTGCCCCAAGNGCTGAAAGAATCAATGGCAGGGGGTCTGGCTGCTTTTC  
ATCTGANTGTGGGATGGGAGGGGNGGTGATGAACATTTNGATATACAATCTACTCTTGAAAATGG  
GACCCCAAGGGGTANCCATCACTTTCACTTATAATTNCCAAAAGAAAGACTACCACCNCTGTCTCCC  
AGAATCNAACNAGTTTCTATTACTTTTNGGNGGGAANAGAATTACAAGTTTTTTAATTNGGAGN  
ACCCTTACAATNCCTAAATCTTTTGGAGGGGGGGTCTTTTATTTAANCCNCCCCCNATCGANNT  
TNTGNTTGGTTTGTNTNNAANCTGTNCTNACTTTGTACNTC

SEQ ID NO: 1014     CGAGGTACTTTGATTCCGTGCTCCTGGCCTTTGGAAACCTGCTGTTCTCTGACG  
GGCCTGTCCCTCATCATTGGCCTGAGGAAGACCTTTTGGTTCTTCTTCCAACGGCACAACTCAAG  
GGAACCAGCTTCCTCCTGGGGGGTGTGGTTATCGTGCTCCTACGCTGGCCCCCTCCTCGGCATGTTT  
CTGGAACCTACGGATTCTTCAGCCTCTTTAAGGGCTTTTTCCCTGTGCGCTTCGGCTCTGGGCAA  
TGTCTGCAACATCCCCCTTCTGGGTGCGCTGTCCGGAGACTTNAAGGCACTAGCTCGATGGTCTGA  
AAAACANAGATGAGCTCCTTGAACCTTGGATCATTTGGTTGAGGGGGCTANGAGGAGAAATGGGGA  
ACCACCNCTCAGNCCCCCTGCACTGACTCNCCTCCCGACATAATCCGGACCTNCCAAGTTCCAAAN  
GAAGGAATGAACTGAGCAACTGACGTCAANTCCCAAANTNACTTANGANGCTGCCAAGAAACNA



ANATTCCAAACCCAAGGAGACTGGCCTGGGCTGGNATTCACACNCTCACTTTTTTTTATNGGANGG  
AAAAGNGANANTAATTNCCAATTGGGCGTGGGGAAAAA

SEQ ID NO: 1015 GGTACAGCAGAGACCTTCCTGCTTTTACTGGGGACTCCAGATTTTCCCCAAA  
CTTGCTTCTGTTGAGATTTTCCCTCACCTTGCCTCTCAGGCACAATAAATATAGTTATACCACTGC  
CCATCAAA

SEQ ID NO: 1016 ACTTTCTGTTCTTTGGCACATTTTGCCCAGCGGATGCAACTTCTATCCTCAGTC  
CAGTTCATATATCTCAGGCAGTGATCCATTTTGTATCAGCCAGTTTCCCTTGTTAGGGCCGCTACCC  
GGCTGGCAGCCAAGAGCAGCACTCTCCCCACCTTCAGCAGAACTTCTCAGCTCATGTGTGTGTTTT  
TAAAAGCAGTAGGCAGCTTCCATCAGCAATGGAAGTGTCCCCCACTCCAGTGGTGGAGGTGCAGT  
CGCTGGACCTGAGGAAGTGGGGAAGCATTACAAACCCCTCCTGCAGTAAGAATGGCAGAGTTGGGA  
AGGATCTTCTCTCGCAGGCTGTTGGCCCACTCCTCTTTTCTGTGTATGAGGGGCATCTTAGTTTTT  
TATGGCGTTTTCAAAGCAGCATATGAGGATTCCTTANCATGCAGCTAANTGACAAACCCCATGGCCT  
ATGATGNAATGTTTGCTCTGGACTAACCNTCTTGATAAATAGGAGNGNATGTATCTTANGGTCTTT  
GANAACACTNTTTTTTTTTTTNNTTAAAANGCCTNTGAAANCCTTNAAANGGCCTNNNAAGGNA  
NCNNNGNAANGGANGGGNCTGG

SEQ ID NO: 1017 GGTACTCAAAGACGAATCATGAAAAAGAAAAAAACTTTATTTCAAACAGGT  
TCAGTGATATATGTGTGTGCTACAGCAAAGGCTGGTTGTGGCAAAGTTTCATTTCAAACGTATGA  
TGTGGGCTGGGCAAGGTGGCTCACGCCTGTAATCCCAGCACTTTGTGAGGCCGAGGTGGGCTGAT  
CACCTGAGGTCAGGAGAGACCGGCCTAGCCAACATGCTGAAACCCCGTCTCTACTAATAATACA  
AAAATTAGCCAAGTGTGGTGGCGCGCACCTGCAATCTCAGCTACTCGGGAGGCTGAGGCAGGAGA  
ATCGCTTGAACCCGGGGGGCAGAGGTTGCAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAG  
CCAGACTCAAAAACAAAACAAAATAAAACAAACAAAAAAACAGAACTGCATGATGTATAATTTT  
GACATTATGTGGGAATGTTAACTTCTGCCAAATGTAGATTCAATCCAACATTATGCCAATTTT  
ATATTAATTTTAGTCCTAAGTTTCATAACCCAAAAAAGGAATACCCCTACCGT  
TCCTCTAAGGGACGAACTGGGAATAATTCCCAAATTG

SEQ ID NO: 1018 ACTTTTTTTTTTTTTTTTTTTTTTAAATCAATATTTATTGGGCGCCTATTATGTG  
CAAGGCACTACACTAGGCGCTGGGGAAGATACAAAGATAAATCTGACAGACTGCCCTCAAAGAG  
CTTACAGTCTAGTATAGGAGCATAACAGTCTCTGGANAANATATTTTAAGTGTAACCTCCCCC  
ATCCCACCCCAACAAAAAAGAAAAACCTACTANACTTGGTTTCTTCCATTTACTTTAGTTTA  
GCAGCTTACGTAAAAAGCATAAATCTGAAAGTCTTTTAAATGCATACTTTTACTGGTAAACAAA  
ATTCATTTTCATTAGAAAAATGCTGAATATTTATTGCAATTAANAAAAATCTTCAGCCGGGCACG  
GTGGCTCACACCTGTTATCCTAACACTTTGGGAGGCCGAGGCAGGTGGATCACTTGAGGTCAGGA  
ATTTGANACCAGCCTGCCAACATANCAAAACCCCGTCTNNCTGAAAAATACAAAATTACCGGACG  
TGGTGGCGGGCGCCTGTAATCCCANCTACTCCGGAGCTGANGCAGGAAAAATCGCTTGACCTGGGG  
GGNNGGAGGTTCAATAAGCCAAATCGCCCCCTTCTTTTACCCTGGGGCTAANACCAAATCTCC  
GNTTCAAAAAATANA

SEQ ID NO: 1019 ACATATATCAATCTCCCTTGCTTGTCTTTAAGAAAGGGCCGTTTCATAGCATT  
GGCACAAACCCTCTATTTCTGTTGCATTTGCATGATTTTAAATAAGAAGGAAATAAACATTTGAT  
TTATTTTCATGCTTCCTAAGTTTCTGGGCAGGGACATGCCTTACTCTTTAGAAACCAATTCCAAGT  
GACATCTGACTGCATTTTCTGTTGGTCCGAACCTCTAAACAAACACTCATAAAGTAAAGTTAAAC  
AATTTGGAGATGTATGAGGAAAAAGTCTTGTCTGTTCAAGTTCAGACTTTGTTAAAAAATAA  
AANNNAANGGAAAAANTGCTCATTTTCACATGTCCATGATCTTCCATGGATTTTTTTTAAGCTTA  
TTTGAAGTTTGATTAAAGGGACAAAAANAANAGGCGGCAAGTTTCTCTATCTTTGGAGNGTTT  
CGCTCAGGAAATTTTGCTCATCAAAGTTCANCTACATTCCNAGCGGACAATNAAGGCAAACTGGG  
GTGCNCC

SEQ ID NO: 1020 ACTTANCTGTATTTTCAAATAAGTAATCTTCCCCCTTTTGTAGGACTTTAA  
ACTAGGCATCAATGAACCTGTTTTTCTTATTATGCCTGGAATNTAGTCATGATATCATGATACCTTG  
ACTCATTCCATCATATTTCAAGAGGATTNAGAGTGCTANAAATTATTTGGTATCCTGTAACACAC  
GGCAACACTGGTCTTGGGCCTATNATGACCCACAGATGACTCATTATAGAGTTCATTGCTGATTT  
NTAANTTANTAATTGAATCTTTATGATA

SEQ ID NO: 1021 ACAAACACCATACTTGGGGCTATATGCGATTTTCAGGTTGGATAAACGAGTC  
ATGTTGAATGACAAAAAGTTAGACTGGGGAGATTTATGGAGAGAAGCAAGCACCTGTGATTTGTT  
GGTTTGGCGTTACAATAAAACACTGTAAGTGAAAATGAGTCACAGGCAGTTAGAATGGCGTTAAGG  
AAACCATTTAGGTCTAAAAATCAGGACTGAAGTGATCAACAGTTAACAGATGAAGCCATGAGAAC  
TGGCACNCTCTGAAAAGGCTGGTGCANAGAGCAGGACGGTGAGCCTGANTGAGCCTGGAGTTAGT

CAACAGATTCTCAACAGAAAAGGANTGAGGAGCANGGTCAAAGAAATGAAGGTGAATTTGGGG  
GGTGGT

SEQ ID NO: 1022 ACATTCCACAAGCATTGCCTTCTTATTTTACTTCTTTTAGCTGTTTAACTTTGT  
AAGATGCAAAGAGGTTGGATCAAGTTTAAATGACTGTGCTGCCCCCTTTCACATCAAAGAACTACT  
GACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGCANGGAAGGAAAGAA  
CTTGCATGTTGGTGAAGGAAGAAGTGGGGTGAAGAAGTGGGGTGGGCCACTGTGAAANTCTAG  
ATATAAACCA

SEQ ID NO: 1023 ACTACACCACTTTTCTCACCACCCCATCCTATTCTTGAGTTGCAGGATAC  
ACTTGCTCTCTGGAAGTGTGTCTTACCCTTCTGCAGAGTGAGGAGCAAGCTGTTAGAGATGCAGC  
CACGGAACCGTGACAACCTGCCATGTCACAAGAAAATACCTGCCAGTCAACAGAGTTTGCCTTCT  
GCCAGGTGGATGCCTCCATCGCTCTGGCCCTGGCCCTGGCCGTCCTGTGTGATCTGCTCCAGCAGT  
GGGACCAGTTGGCCCTGGACTGCCATCCTGCTGGGATGGCTGTTGGGAGAGAGTGATGACCTC  
TTGCTGTGTGGAGAGCATGCATCAGGTGAAGAAGACTACCTGTTNAAAAA

SEQ ID NO: 1024 ACAAATTTGCCACAGGTTGAACACTTAATTTGTGTTTCTTAAAAATAATGCT  
GTGCTTAGTTTATTGCCAGGAACTTCTTGCTTGATGCTTTGGTCTGTTTGTITTTAATGTTTTAAA  
CAAAGAATATTCGGGCTTTACTTCAGAGTTTCTTCCAAAGATACATATGCATTTTGGATCACAGT  
TCTGGAGGTTGAAATGTCCACGATCAGAGTTCCTGCATGGTCTGGTTCTCGTGACGACTCTTCCTG  
GCTGACAGATTGGGGCTTGCTGTATCCTCAGTGGCATAAAGAGAGCTCTTGTCTCTTCATTCCCT  
TGNAAGGCTCCAATTCCATCATGAGTTCCCCCGCGT

SEQ ID NO: 1025 GTACTCTATTTCGTATTANGAAAGAGAGGCTAGATACCAAACATCACGAGATC  
ATANAGATGGAANACTGTATANTAACCAAAAACCAAAACAGTGATATTTCCAGAACTTTNTAGNA  
TNGTTGTAAGAACTGCATGAAACANAAGACCATTTAANTAGGTGAATTATTANAAATNTTTATGAT  
GGGAAAGNNNANCTAAATTNTCTCTAAAATTTTCCCCAAAGATAATNNCCTCTTAGGGGAACATA  
NGAGAAATGNTTNATCAAANGTGGGGTNTTANTTNACCNTGACTATGTTTTNAGTTGCCNAGTCN  
TAGT

SEQ ID NO: 1026 ACGCGGGGTTCTCGGCTGGATTTAAGGTTGCCGCTAGCCGCTGGGAATTTA  
AGGGACCCACACTACCTTCCCGAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTCA  
TGGGACCTGTGCGGTTGGGAATATTGCTTTTCTTTTGGCCGTGCACGAGGCTTGGGCTGGGA  
TGTTGAAGGAGGAGGACGATGACACAGAACGCTTGCCAGCAAATGCGAAGGTATTTGAAGGGG  
GTAGCCCCTATAGGCATCGCCCGGCCACACCTCCTTCTTAGGCCGGACATCCTAGCCTGTATTG  
GGTGGGTACCCGACCTCCTTGAAGGCTTGAGGGAGGTTCTGGTGGAGCTTGGCTGAATTTGGATG  
GGGTCAAGGAACGAACGTTAGTGCCACCTCAAACTGCATCCAGAGAGAGCCTGCATTTTGTG  
GCTGGCTTGGAACCCCCCGCCCCAATGGATTGTGGGAATTGTAGTTCANGGACCATCGGCCCGG  
GACTGCCAACTTTTANAGCATCTTCAGTGAGTGTTTACCTCATGTCACCAGACCTTAATCTGCANA  
AGCTTGCAAGAACCGGGTGTGANGCTCAAGCCTTGNGGACATTCGGTTGCCTTACCTACATTANAC  
CCTCCCAA

SEQ ID NO: 1027 ACATACACAAAAAGTTACTGGAATGCTCGGAATAAGATTGTTTTTCTGTTGT  
CATTTTTGCTTTTTTTACAAGGTTTTTTTTCTCCTTTGAGATTATAATGAACATGGTCACACCACAA  
GTAAAGTCAGAAAGTAGGACAGAGAACGCTCCGAAGGCTGGTTTGGTCATCCGAGATCATTAAGAAA  
TGGCTGACCCTAACAAATATGTACTTTTTTTTTTTTTTTTTTTTTTTGGGTGGATTGAACANAAT  
TTATTGGCTGTCTTTGAGTGTCTTTGGTATGGCTTTGGCAGGGCTGTCTGGGTTCCTCCGCTTTGCT  
TGTTTTTGGGCTGTCTGCTGCAGCCTTAAGGCTCTTCTTCGCTTCTTCAGCTTTTGAGTCTCCTGGA  
AAACCCGAATGCACANAGCCTTCTTGGTCACNAATCGGCGNGGGCTCTTCGGTAATACAGAGGG  
GGGAAGGTGTNCACTGTTCCAAAAGATCTCAATAAATTACTGGCTTTGGTAGGTGGAATANGTTCT  
TTGGGTGGGACTTGNGGAGCTNCTGGTTTAAAAGGNTTCTGTCNANTGTT

SEQ ID NO: 1028 ACAGTCCCTCTCCTATAAGCAAGAAGCTCTCGTGTGCTAGTGTCAAAAGCCA  
AGGCAGACCGTCTCCTGCCCTGCTGGGATGGCTGTCACTGGCTGTGCTTGTGGCTATGGCTGTGG  
TTCGTGGGATGTTTCACTGGAAACCACTGCCACTGCCAGTGCAGTGTGGTGGACTGGACCACTGC  
CCGCTGCTGCCACCTGACCTGACAGGGAGGAGGCTGAGAACTCAGTTTTGTGACCATGACAGTAA  
TGAAACCAGGGTCCCAACCAAGAAATCTAACTCAAACATCCCACTTCATTTGTTCCATTCTGATT  
CTTGGGTAATAAAGACAACTTTGT

SEQ ID NO: 1029 ACAGGAAATTGACTTAGCACTTTCCCTGTTTTTCTATTGCATAATTTTTTTTTT  
AACCCAAAGATATTTTTTTTGTGAGCCTGCCAGTATTCAGTGTTCACAACTTTGATTACTGGCTA  
CAAGAAATATTTTCTGCTTCCCCAAATCCCATACTCCCCAGAATCTGCTGGCAAAGTGAGCCCT

GGTACTTTTTTTTTTTTTTTTTTTTTTTTACTTCACAGCAAGATTTATTAATACATCCAAAANAAAG  
 AAAGAAAAATGATTTTGTACACTATNTACAGATAATACATACAGNGTTTCATACACATATTACAT  
 CAGTTTTTACACAAGAAAAATATACATAAAAAATGTAACTTTTGTNTACAAAAACCCTTAAACAA  
 ATTAACAAGGACCAGATAACAAAAGGAGACCAATTTCATTATTATTCATCAAAATTTTAAAAACA  
 GTAATACAGTACCACCTGCTCCCATGAACTCTTAAATCTGAATAATGTTTCANATAATTTTCATAGG  
 NGGAAGTGGCAAAACATACAACCAATTGTCCCTCCTCCCCAAGTCCACATGCCAAAATGTTTGCA  
 TATAAAGTACAAGANGAACCCANAAGGGGACTAAATACCACTATGCTTTTTTCTTCACTAAAAA  
 ATAG

SEQ ID NO: 1030 ACACATGCACCTTTGTGCTTTTACACACACAAAAAGTATACTGTAATCCACTGAG  
 AATAACCTCAGCTGGGTCTGTTTCTGGGTATGTTATATCTTGTAACAAAAACAAAAACAAAA  
 CCAAAAAAAGTTTGGATTGGCTGGGTCTGCTCCCACTTTCTGAATCAAAATGCCCAACTAC  
 GTGGTTCTGCTGCAAGTCCATGAGCAAAGACGACTCAGAGGGGTGGGCAGGTTTGGTATCACCAG  
 AACAGGAGCATCTACCATGGAAACACCACCTTCTGTGGCCCTCATTGTCAGAGGGGCAGAGTT  
 CCCGAGGGATGTGTCCTCTGGGACTGTGTGGTCTTAAAGTAAAGGTATCCTAAAATGGTCAGAA  
 CATGCAATTTCTTTCAAAGGCAGTTCAGTGCATTGGCTCAAGGAGGGTGCAAGCCCAGGATGAA  
 GTGGAGTCTTGGGAGGGCCGCACTCCTGAAGCCACCAAGCAGAGCGAGGAGCTCCGGGGGGTC  
 TTCTCTGCTTCCATCCTGCGTCTCAGTTCCTCCCGACCCTTGCTCTACTTCACAAGACAGANGCAT  
 TCCTTTCCAAACCTTCCATTGAATGTCGCTCTAACATGANCCGCCACGAAATGGACTGGAAANA  
 ACCTTGTG

SEQ ID NO: 1031 ACTGTATAACATCTTGTTTATTATTAAATGTTTTCTAAAAATAAAAAATGTTAGT  
 GGTTCCTCAAATGGCCTAATAAAAAACAATTATTTGTAAATAAAAAACACTGTTAGTAATAATCAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAANGTNCAGTCAATAACACAATAGAAAAATCTATATACAGTGTG  
 TTAATGAGTATGGAGGAAAGGCAATAAATAGGTCATAGTANCGAAGTAATTACAATTTAGTA  
 AATTAACACTGGAATGATGGATGTGCAGATGATGATGTGCAAGTAGAAATACTTGTGTGCAAAAA  
 AAAAAAAAAAAAAAAAAAAGGTCTCAAAGACNANTCNTGAAAAANAAAAANCTTTATTNTN  
 AAACAGGTTNAGTGATNTNTGTGTGTGCTACANCAAAGGCTGGTTGTGGNAAANTTTCATTTCAA  
 ACTGTATGATGTGGGCTGGGCAAGGTGGNTCACGCCTNTATCCAAGCANTTTNNGAGGCCNAGGT  
 NGGGCTGATCACCTGAGGTCAGGANAGACCGGCCTACCAACNTGCTGAAACCCNTTTTACTTA  
 TTATNCAAAAATTANCCAAGTGTGGTGGCCCCCCTGTANTCTCANNTACTCGGAAGCTNAGGCA  
 GGANAATCTNTT

SEQ ID NO: 1032 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNGNGAAAAATACTTATTTTCATGT  
 GTTAAAAATACATTTTATGGGTGGGCCCTGCAGGAGGANACAGGCCGTCCACATNTCCTTCCCAA  
 TAGTGTGTCCAGTCTGTTTCCAAAATCTGTGGGTCCCTCACAGCTTCTGATGACAGCTGCTAATGC  
 CATTTGCTGAGGAACAAGGATGGGGAGGATGGCGAGGGCCTGGCCCCCAGGGCGGCCACACCAA  
 AAGGTGCGANAAAGGCCCAAGGCGGATGCCACNCCCAGCAGTGGTGAAGTCCCCCACTCCTTTT  
 CTGAGTCTATCAGCATTGTTTGGTTTCTTCTTGTGCTTTTAAATCCTCGAGCGGCACCTTGACCA  
 AAGCANATGGTANAGGGTATTCTATGTCTGTAAATCCGCCTNTCACCTGCCGGTGCAAGGGGAGGC  
 TGCCACAGGCTCANAACANATTTCCATGGCCTCCCCTTTTGACCGCTCTCCACTGTCCCCTGGGGC  
 TTCTGCTGAGCTGGGCTTTGCTTTTCTTTGGCTTATNAAACCGGAAGTCCGGGGGTCTNNGNCGN  
 NACCACCTIANGGGCG

SEQ ID NO: 1033 ACGCGGGGAGGCTTGAGGGAAGCATGGAGGTCCATGGCAAGCCCCAGGCTA  
 GCCCCAGTTGTTCTGTCGCCACCCGGGATTCCTCAGGAGTCCAGTGTCCAAGGAGCTGCTGACG  
 GCGGGAAGCGACGGCCGCGGAGGTATATGGGACAGGTTGCTCATCAACTCCCAACCTAAGTCCAG  
 AAAGACCTCCACTCTTCAAACAGTTCCGATAGAGAGGAGTCCCTTATTGGACCAGGT

SEQ ID NO: 1034 ACTGAAGCAGCATATCAATCCCAATAAGACATTGGACCCTTTTGAAACCATG  
 CTGAAGTCATTATTAAGGTATCAATCTGGTGGTGGCAGTGTGAGTGAAAACCATGAGGAAAAA  
 ATTGTATGAAAATGGTGTGACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCTAAAGCTTTCAGA  
 AGAATTATTGATAAATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCTCAGCCAGCATATGCT  
 TGGTTTTGCTATGAAGTCTGTTACACAGATGGTAATGGGTAGTACTTNNTTTTTTTTTTTTTTTTT  
 TTTTTTTTAGTANANACAAAGTCCACCATGTTGTCCAGGCTCCACACTTNTTTGAAATGAANAGAT  
 CAAACATNTTAATACATACTTATTAATGCTTACTTTTTTTTAAACATTTAATATTTATTTATTTT  
 GGGATGGAGTTTCACTCTTGTGCTAGGCTGGAGTGCCAGTGGGCCCCGCGT

SEQ ID NO: 1035 ACTAGGCCTATCAAGAAAAAGCTGTGAACAGCAACATCATAGACACAAAAGG  
 GTCATTCTGGGTGTCCTCACCCAAGAAAAAGATGGAGGCAAGTTAACACAAGATTTTTTTTAAA  
 GATACACTAAAATCCCGCGTACTTTTTTTTTTTTTTTT

SEQ ID NO: 1036 ACTTTTTTTTTTTTTTTTTNTTTTGCTAAAGCGTTTATTTCCATTCAATTCCTAA  
CTGATTGCTCTGCTTGGGTGTCTGTGTTGGTTCACAGGAGGGAGTCTGTGGCCAGATCCCTAGACT  
GAGGGCGTCCCCATGGCCAAGACCTTTAGGCCAGCAAAGGTCTGGTGGGGGAAGGAGGTGG  
ANA

SEQ ID NO: 1037 ACCAGCTTCTAAAGCCATGATGCCATAGGTCCATTGTTGATGAAATTCCTAC  
CCACTGTCCTCGGGCATCTGACTCTGGTCTCTGCACTGGCATCAAGAGAACGCTGCTCGGTGGTTT  
AAGGCTAACACCTTACAGGGTAACACTGTAACACTGGCCCTGGAGCCAGGTGCTTTTCTCCATGAA  
AACTTCCACCTTGGTAGCTCAGCCGACATAGACAACACACAAAGCGCAGCTCCGCACTTCTGTCT  
TATCTTACACAGTGACATCCACACCAGGTGGCCAAACAGAAGAGAAGGCAGAGGCCACCAAG  
AGCTGATGCTGCGCAGTCTTGGGGGATCATCCTCCGGTCTCACTGGGGACGAACCCAGGTTCTGG  
AGCCTCTCCCCTGACAGACAGCTTGTACCCGGCACTTATGGGTCTCTGGGATTTCAGACAATACC  
CAACTTCTGTAGGTTTCAAGAAAGTGCTTTCAAGCAGGCAAGTGGCACCCACACCCGGTGGGACACA  
CCTCTGGGTCCGAAACCACTCCATCATGTGCGTTGGTGTGGCCACCGTGCCACAGGTCAGGCAG  
AAATTGGACGATCCCCCGCACAGCCACGTTGACAGATGGGCACACTGGAACGNTGAAAGCCTTTG  
CANATGGCACAAAC

SEQ ID NO: 1038 ACNCGGGGGTAATATGGTNNAAGANAACCCATATAGACAGCTGCCTTGTAAC  
TGTCATGGAAGCATGCCTGGAAAGACAGCAATNGAACTTGGACCTCTGTGGTCAANTNCCCTTA  
CAATACTGGATTCTCAAAGAATGCTATTNGAATCTNTTCACCATNGGNTTGGATGACATTGAG  
ACCCTATTAANGACNTTNTATCTNTGAANTCNGAGTGACGCGGGGCCTTCTCCCCG

SEQ ID NO: 1039 ACTTTTAAATTTTTTTTTTTTTTTTTTTGGATTTTTAGTAGAGACGGAGTTTTAC  
CTTGTGGCCAGGCTAGTCTCGAACTCGTGACCTCAGGTGATCCGTCTGCCTCAGCATCCCAAAGT  
GCTGGGATTACAGGCGTTAGCCACCATGCAGCCCTTTCAAGCCTTTTAACATCATGTACCTTCA  
CAATGAGCAGTTGCTCCATTACCCAGTGAACCTCCTGATGGGACAGTTACCCAGTTGGGG  
CTTCTCAGCCTTGAGAGATGTGCAGAGCAGGGACCTGTCCAGGGCAGAGCAGCCAGACAGCGTGG  
AAATAATCCAAACAGAGAAGGCATTCAAGAACTTGGCCTCTGGTTGACTTAACACACTACTTCAA  
GTATATGGCAGGTCTGTATAATAAACTCCTGCCTGTCTAGTCTCAGCCTCATAGAAAGCAAGAA  
GCAGGTGGGCGTATGTTGCAAAACAATTCTTGGGTAGAACAAAGGGGGACATTTGATGGGTTTGGG  
TTTTTTCAGTAANAAAGATGAAAGTCTGGAAGTTTTAGGGTAAAGTGGGTACTTTTTTTTTTTTT  
TTTTTTTTTAAAACTCNNGGAAGGTNTTTTTTNTAGGGATTGATNCATGGGGCCNCATTTGNATAA  
NTGG

SEQ ID NO: 1040 ACAGTAATCCTGTGAGAAAAGACAGGACAGAAACCACTGTGCCTATTTTACAG  
ATACGAAAACCTGAGACACAGGTAAAGGGGCTTGTCTGTAGTCCCATAGCTAGCAGATGGCTGGAG  
CCAAGACTGAGGCTCGTTCTTCAATGCTGAGCCAGGGCTCCTTCCGCTGCACCACAAGAACGCTA  
GACCACTGCCACAGCCTTCTCATTCCCTCTTCCCTCATTCTAATCATTTCTAGCTGGCTGGCCTC  
CACAGAGCATAGGAAAACAGCCAAGGCCGGGCACGGTGGCTCATGCCTGTAATCTCAACACTCTG  
GGAGGCCGAGCCGGGTGGATAACCTGAGGTCAGGAATTCGAGACCAGCCTGCCAACATGGTAAA  
ACCCCATCTCTACTAAAAATATAAAAAATTAGCCAGGCATGGTGGCGCACACCTGTAATCCCAGCT  
ACTCAAGANGCTTGAGGCAGGANAATTGCTTAAATCTGGGAGGCGGAAGTTGCAGTGAGCCAN  
ATCGCGCACTTGAACCTCCAGCCTTAGGCACAAGAGCAAAAACCTCCATTCTNCAAAAAAAGA  
AAGGGGAAAAACCAGGGGCCAGGTNACCCATTTGGGGGAAAAGAAGCCCCACTTTAGGAAATCC  
TGGGGATTGTTAAGTGT

SEQ ID NO: 1041 ACTACTAGTGGACTCAAGTGATATAAAAAAATAAAAAATAAAATACTTCCAT  
ATACAGTGACTGGACTCTTAATACAATGGAAAAAATACTTTGGTGAAACACTGACATTTAGTGAG  
CCCCTATCGTGAGCCTTTGGGGTTGGAAAGGGGAGACGGGGAGTGAGTGGAAAGATAACTAAGA  
CTCAGACCTCATTACAGAGTCATTCACAATCCATGTTCCACGTAAGAAGGCACACCTTTGTGTAGC  
TGTGCACAGTTAATAATAGGGGCAAGATGTGTTATAGCAGAATTACCAAATGCTATAAAAGTGGC  
ATGAAACAAAGAAGTTCTTCTAATTGGGGTAACGGGAGTCAAACCTTACAGAAATTTGGGAGAGAGC  
ANTAACATTTAGCTTGAATATTTTAATTTTAAACACATTTAAAGCTTCTAGACTCACANGTACCAG  
GTGCTCAAAGATGTTTGANGGTTGTCAAGCAAGGCTGTTCAACTTCTGCATTTTCTGTGGATATGG  
GCACAATGCCCACTCAGATTTTCTNCTANGTCTGAATGACAGCTTATCTACNCANAATGGAACT  
CCANATTTGTGACANCACATGGACANATGCTGCAATCANAACAATTCTGGATGTAAACACTGGTA  
GGCTGCACTGGTC

SEQ ID NO: 1042 ACGCGAATTGGAGAAAAAGTTCAGTGGGAAGCATGTGCTCTTTATCGCTCAG  
AGGAGAATTCTGCCTAAGCCAACTCGAAAAAAGGTTNCTTCAA  
TGTGATTGTAACAATACTCTGATCAAGGTGAGGTGTGTTCTTGGCACATGCTGAAGTTTTCTGTC  
AACTTATTCCAGAGAGACTATCAAAGGCATGAAGATCTAATACATTAGAGAGATCAACTCTTGA

TCTTTGTGTTCTAANATTTGTCCTAGTCCATTTATGGATCTAATTGTGCGGTCTTAATINCTTTACATC  
 TTTCTTCTGGAACCAAGGTCTGTTTTATTAATGGANAATGGCATCTGCCAAAGCAACTTGCCTAGT  
 ANCTTTTCATTGATTAAGGCCATCANGGTTCTNNTCTGGTTAAATGGTTTTAATCCAATNNTTTNGA  
 ATNCCACCAATAGGTTANTNAATNCCCNTCAAAGAATTAATAATTTTCACTTCCCCTAAATTTNAGN  
 CATTNAACCCCAAAACCATTANNAANCCCCCTTGCANCCCAGGGGGTCTNNTGGCTTAAATTCC  
 CAGGGGGGGNCCTTTTTAAAACCANNTTTTTNGTNAAATTCCAAAAATTTTNNCCCCNTTNTTTTTT  
 GGCCAT

SEQ ID NO: 1043 ACGAGCGGCGCGCCATGGAGTTACTGAAGGTCTCCAAGGACAAACGGGGCCCT  
 CAAATTTATCAAGAAAAGGTGGGGACGCACATCCGCGCCAAGAGGAAGCGGGAGGAGCTGAGC  
 AACGTACTTGCACAGGAAGTGTTGGCGCTTGTTCATTCGTTGCTGCTCCAAGTTAAAAAGTTGTT  
 ATTGGAGCTCATCTCAGCACAGTGCTTGTTCACCCATGGACTTGCCAGACCAGGATCTGT

SEQ ID NO: 1044 ACTTTCTGTTCTTTGGCACATTTTGCCAGCGGATGCAACTTCTATCCTCAGTC  
 CAGTTCATATATCTCAGGCAGTGATCCATTTTGTATCAGCCAGTTTCCCTTGTTAGGGCCGCTAGAC  
 CCGGCTGGCAGCCAAGAGCAGCACTCTCCCCACCTTCAGCAGAACTTCTCAGCTCATGTGTGTGTT  
 TTTAAAAGCAGTAGGCAGCTTCCATCAGCAATGGAAGTGTCCTCCACTCCAGTGGTGGAGGTGCA  
 GTCGCTGGACCTGAGGAAGTGGGGAAGCATTACAAACCCCTCCTGCAGTAAGAATGGCAGAGTTGG  
 GAAGGATCTTCTCAGCAGGCTGTTGGCCCACTCCTCTTTTCTGTGTATGAGGGCATCTTAGTTTT  
 TTATGGCGTTTTTCAAAGCAGCATATGAGGATTCTTAGCATGCAGTAAGTGACAAACCCCATGGCC  
 TATGATGTAATGTTTGTCTGGACTAGCCATCTTGATAAATAGGAGTGATGTATCTTAGGTTCTTT  
 GTAGAACAGCTATTTTTTTTTTCTCATTAATAAAGGCCTCTGAAAGCCCTCTNAAAATGGACCTCAC  
 AAAGGCAACCAGTGTAATGGATGGGAGCTGGGGTGCCTGCATGCCAGGCTGGTCNTGGGGTGG  
 GCAC

SEQ ID NO: 1045 ACAGCTTTCTTCGTCTCCATGCTAAGAGATGTAAAAGCTTAAGGGTCAAAC  
 AATACCAATTGTATAGGCTTCAAAAACCATCTAAGTTAGGGCATTCTCTAGTTTTAGCTAAGATAC  
 ACCTGGAACACTGACAAGTCATCACTACATAGAAATGTGAAGTAAATTTTTGAAAAATAAA  
 TTTAGTGGAACAACTCTGAAGGATAACACCAGAGAAGAATAGCAGGTTACCAGTAAGGTGTCAGCC  
 AATTTGTTCCAGTCACTTTTGAATCCATGTTCTATAATCTAAAATTTATTCTCTTTCCCTAAGCTGA  
 GAGCTTCTATCATGTGAGTATCTATGTTATGAAGAAAAGGAGACTTAGGTGAGATGTTTTTATTT  
 ATCGCAACTGCTGCATTAATGCCTAGGACCTCAACAGCTTCATGAAAGTCTGGGAAATGTTTCATG  
 CATAAGGTTATTGCCCCGCT

SEQ ID NO: 1046 ACCAGTAAAACTTAAAGGCACAAATTCTCCTTGAAGACCTTCTCCCTTTTAT  
 GTGGCCCCATATTTTATGTTGCTTTATCTTTGAAATTTGTCATGAAAAGGAAATGAATGGATTCTGA  
 ATGAAATTGTCTTTAGAGCATGATTACTTGTTCCTATGGACAAATATTTTCTCCCTTGCTCTTC  
 CTGGCCTGAAACACGGGAAACCAGAGTCAAAAGTTATCTCCCTCTCCCTGTGATGCCTTGAGATTT  
 TTTTCTGCGTTGTTAATGCCTGAAATCCAAGTCTTCTCCATGGGAAAATACTGTTATAACCAATA  
 ATTCTAGATGAGTAACAAAGATCTTTTAGGCCTTCATTTTATGTTTTTCTTAACTGTTATATTATG  
 ATTGTGACATAGATTATACTACTACTAATTTTTGGATGTTTCAAAGGTCAAGAAGTAAAGATGT  
 TAGAAAGCAATGAGTGAGTCCTTTTGATTTTAACTTATTCCCCATGTCCCTATACTTCGTGTGCTT  
 TTCTTTTTTTTTTTTGGAGACGGAGGCTCACTCCGTACCTAGGCTAAAGT

SEQ ID NO: 1047 ACCCCACAGCTCCACACTGTCATCCCCCAGCCAAGGGCCATCCCTAGAAAA  
 ACTGGTTTACTGTTCTTAAGGAAACCATTGTCTATAGCCCTTAGCCTTGTTGGTTTATGCAAGAGGC  
 CAGGTGAGGTGCCCTANAGTGAGGCAGGGGCTTCAGCCAAAGTTGTGATCGCAGCTTCTGAGGCA  
 GTTCTNAGTGAGTCAAGTCCGCTGCCACCTGAGCTTTCACCAGATCTTCTTTCTTTCTTGT  
 GCTTCTCAAGCTGCAGGCTTGATCCCATCCACAAATATGAGAGAATTCTGGAAAGTGCCCTGA  
 GAAATGGGTTCTGGGTTTTTTTTTCAATTTTATTTTTTAAATGTCCNAGGCTGCCTTNTGTCTCC  
 NATCCTGCATCCTGCTTTCTGTTCCCTGGATACCGTAGGATGGTTTTATTTCAAGTTATGCACAAAT  
 TAGTCTGGACACTGTGGAGTNATAACAAGAGTGGGATGGAGGTTCCAGGGCCAATCANTTTCTTT  
 GGAGGGAAGCTTGAGAAGTAGGTTAAAGCAATGCCAAAAAAGCCACTGCTTNNCTTNCATCCC  
 CACAATGAACTAANACCAACCTAACAGAAAGAGGGTCTTTGATCAANAGGGTCAANCAGTGG  
 GATGTGATTAGTGGGTTNGGGGATGGAGTNCCAGGTAAACCAAAAGG

SEQ ID NO: 1048 ACGCGGGGGGCAACGAGGAGGGCTGCGAGGCCATCAGCTTCTCCTGTCCCT  
 CATCGACAGGCTGGTCTCTACTGCGGGAGCCGGCTGGGCAAATACTACGTCAAGGAGAGGTCTA  
 AGGCAATGGTGGCTTGCTATCCGGGAAATGGAACAGGTTATGTTCCGCACGTGGACAACCCCAAC  
 GGTGATGGTCGTCATCACCTGCATCTACTATCTGAACAAGATTGGGATGCCAAGCTACATGGT  
 GGGATCCTGCGGATATTTCCAGAGGGGAAATCATTTAGCATAGCAGATGTGGAGGCCATTTTTGACAG  
 ACTCCTGTTCTTCTGGTCAGATCGTAGGAACCCACACGAAGTGCAGCCCTCTTACGCAACCAGATA

TGCTATGACTGTCTGGT

SEQ ID NO: 1049 ACAGATGTGTATGGGAAACCCCAACCCCTATATATTGTAAATAGATGGGCTG  
GGCTAAACATTGTTGCCGTTTCATACTTCTACCAACTCAGCTTTTACACAATAAAGCTCTACTGTCT  
CTGGATAAAAAAAAAAAAAAAAAAAAAAGTACGCGGGGATCCCTGACTCGGGGTCGCCCTTTGGA  
GCANAGAGGAGGCAATGGCCACCATGGAGAACAAGGTGATCTGCGCCCTGGTCTGGTGTCCATG  
CTGGCCCTCGGCACCCTGGCCGAGGCCCAAACAGAGACGT

SEQ ID NO: 1050 ACTACTTGATTGTTTATATCCAATTCCTTTCCATCATTGATTGAATCTTTTGG  
GTCTCTTGCTTTTGTTCCTCTAAAACTGGTCTGCTAACTTTTAAATATTTTCTTCATGCTGTGCTCT  
CAATTCCTTCATCTGCTGTCCACACTGAAAACCTTCATATTCTAACCTCTTCACAAGGTAAGTATTG  
TTCTTCCTATAGTCTGAAGCTGGTCTTCTGTGCGAAGAAATTCTTGACGAAGCTCAGCAAGTTGCT  
CCTTTAAATGATGTATGTCTGCCATCTGATACGATATGTTGTTCTGTAGTTTAACTTGTATCCTC  
GCATCTCTTCCCGAGGCCCTCTCTGGCCTGCAGCCGGCTGCTGAGGCGGCCGTAGTCGGCCTCCTT  
CTGGTTCGATCTGTTCTTCCCGCGT

SEQ ID NO: 1051 ACTCTGTTGTAATGGGAAAACATTAATATCTGCTTCTTCTGACACGACAGTAA  
AAGTATGGAATGCACACAAGGGATTTTGCATGTCAACATTAAGGACACATAAAGATTACGTAAAG  
GCCTTAGCATATGCCAAGGATAAAGAACTAGTAGCATCAGCTGGGTTGGACAGACAAATATTCTT  
TTGGGATGTGAATACTCTAACAGCATTGACTGCCTCAAATAACACTGTCACAACCTCTTCTTTAAG  
TGGAACAAAGATTCCATTTATAGCCTGGCCATGAATCAACTGGGAACAATCATTGTATCAGGGT  
CCACTGAAAAGGTGTTACGGGTATGGGATCCAAGAACATGTGCAAACTAATGAAGCTTAAAGGG  
CACACGGATAATGTGAAGGCATTGCTATTAAACAGAGATGGCACGCAATGCCTGTCAGGCAGTTC  
TGATGGGACAATTCGCCTTTGGTCCCTTGCCAGCAGAGATGTATAGCAACATACCGAGTCCATGA  
TGAAGGTGTTTGGGCGCTGCAAGTCAATGATGCCTTCACACATGTGTATTCTGGTGAAGGGACA  
GGAAGATTTATTGT

SEQ ID NO: 1052 ACAAAGAATACAATATGATTTGTCAAAAAACATATAAAAAAGACAGCTGCTC  
TTCCTCAAATACATGAGCTAATGATAAAAGACTTTTCATGTTAATGTCTCCAAGTCTTCTTTTA  
CATAAAAAAGAACATTATGGTGGCAAATGTGAATTATCCTTTTAAATATTGAACATTATATTCTTTT  
AAAATCCATCCAGATCAAATGCAATAATTTTCTTTTAACTCAACAACTGATGCTACCAAACGTGG  
ACTCAATATACTTGTTAAAACGTGTAAAGCGTGTCTCTAGTCTTCAAAGCTTTTCAGGTGAAGAGAG  
GTGCTTTTCTTGATGCAAATCTCAAGGCAGAGAAAATCATTTTAAAGCTTATAAAAAAGTGGACAG  
AGAAATATTAAAACTTCTCTGAAATATACAAATATGTGTAATTATTAATAATTGAAGACAGTAAC  
ATCAGTTGCAAGTGCTTGGAAGTCTGCCTGACCTTTTGAGTTTCCACATTTTCTTCAGTTGCATGGT  
GAGCACCCAGTGCTAGTAAAACATTTCAGGCTGGGAAAGGAGATATAGNCTTCATTACTGNTTTT  
AAGAAAGAAAATTCTATGNTTTTGGCAGTCTTCAGTATCAAAGTATAGGNNTTCNATTACTAGAA  
ATTCCA

SEQ ID NO: 1053 ACTTTTTTTTTTTTTTTTTTNGGGGNNNCCATNAAAAAGCTTTATTTCCATTTGG  
TCCAAGGCTTGTTAGGATAGTTAAAAAAGCTGCCTATTGGCTGGAGGGANAGGCTTAGGCAAAAC  
CCCTATTACTTTGCAAGGGGCCCTTCAAANTCNCCTGGGCTCAAAGGCTCTTAGTCGTGCTTGAA  
AGNG

SEQ ID NO: 1054 ACTTTTTTTTTTTTTTTTTTTTTTTTTCATTCTACTTTTCTTTATTGTCTGGCTAAC  
TTACAAANATGCANATGTCTAGGGTAGTCTNTACCCTACCCTTACACTATCCTGATGACACANAT  
AGCAAAATGTGTCTGTTTACATAGTGCATGATGAAAAAAAGTTTTTCTTCTCTACGGTCCCT  
GACTATAAGGAGGGAAAAATTAATTTTCATGCCAACATTTTGGGGAACTTTAACAATCATCCCATT  
TCTGCTACTAAAAATAACAAAACCTGGTATTACACTTTAAAAATATAAAGACCTAACAGTTTTTACAAA  
TATGCAAATAATCTACTACTTAGACATAAAAAAAGTTGATTTCTTTTAAATCACAAGTAAGGCC  
CATTGGATTAAACATTCTCTGCTTTTACTAAATAAAATGCATAGTGAAATAAATACTGAACACTG  
AGTTTAATACTGNATACATTCAATATAAAATAAGAGGNGAATGTAAANNCTGNTACTGTGNATCA  
TTATCTGAAATGTNTAAAAAACNNNTGTAGCTTGNTTAGGAAAAAATATTTTATTTCATATTAAAT  
TTNC

SEQ ID NO: 1055 ACCAGAATACTCTCACATTTTATTTCAAGGCACTCTCCAGCAGAAAGTATAAT  
AAGAATAAACTACTTTGTAAAAATGTGCCATTTTATAAAGACAACAGTGAGATCCAAAATGAGAC  
AAAGTAGAATTATACTGAATGTTTCATTATCCCAAAGTGGCTTAATTTTCATTGCTTGCTGCAAGA  
CAAAGACTGTAAAAAGGAGCAAAGAAAGTGTGCAAAGCACTTCCTTGGAAGACCTCGTCCACTAA  
CAGATCACAAGAAAAGCAAGTGCAGCCCCTCCACCGAAAGGAAGGCAGGCATGTAACATCGGAG  
ACTTCCAGTCTCCCCGACGTCTGCTCACGCTGGATGCTGGCACTGGCTTATCATGATGGAAGGAGA

SEQ ID NO: 1056 ACCTCTCATTTTGGCTCTGCCCCTCTGGAAGATCCTCTCTGTATCCTCATCCTG  
CGATGAGTGGTGGCAACGTGTGCCCTGAGCCCTATGCTAACGTGAGTGGTTTCTTTCAGTGTTC  
AGATTTTCCCCAGCTCAGTCCCTCCCTCTTTCTGCAGCTTGGTCTGGTTTCTTCTTGCCGGCGTCT  
CCAACAAGCAATGATGGCTTCATCGTCCATTTCTTCTCTTCTCTCTGTCTCAGTACTTTTT

SEQ ID NO: 1057 ACGCGGGGCACCGTGGAGAGCAGAGCGCGGCGGCTGGAAGCTGCTAAGTCA  
GAGCCGCGATGTTCCGGATTGAGGGCCTCGCGCCGAAGCTGGACCCGGAGGAGATGAAACGGAA  
GATGCNCGAGGATGTGATCTCCTCCATACGGAACCTTTCTCATCTACGTGGCCCTCCTGCGAGTCAC  
TCCATTTATNTTAAAGAAATTGGACAGCNTATGAANACAGGACATCACATATGANTGCACGATAT  
TAAGAGCCTGGTTCAGTTTCTACTCCTCTT

SEQ ID NO: 1058 ACGCGGGACTCTAAAACTGGTGGTTTCTTACTGAAGGTGTTCTCCATTTGAAA  
TTTTATCTTCAAAGTATTTATAAGTANTATCTTTAAGACATGACTTGTTAGTAATAAAAGTGTTACT  
AGNCGAAGAGTANCTCTCAAATTTCTCTTAATGTAAATCACCTGGGAATCNCCTCAAGTNTTTTT  
GAAATTTTAAACCCAC

SEQ ID NO: 1059 ACAGAGAAACCACAGGTTGCCCTTTCCACAGCTGGATAGACTTATCCAAAAC  
GGCAGGATGGTTCTGTATTAATCTTTTTGGAAAGCATGTCTGTATTAAGATTGCAAAACATACAGA  
TAGCTACCACAAATTAGGTCAAACGACTGATCAAGTTGTAACATCTGTGAGGTCAAATTCAGTTA  
TAATAAAGTGCCTAGATACACATTTATACAACAGACCATAAGAGCTGAATTTTACAAATGTCTT  
TATGGGCATGTAAAAATTGACTCTGCATTTCTGCATGTGTGCATTACATAAGAGAGACCAGTCTGC  
ACTGAGTCATATATACTCCAACTTGAAAAAGTAAGTGTAACAACTGGTTAATCATGCAAGTCTGTT  
TGTAATATAACAATGACTGGTAAACATGAATTTCTGCACAGTAGTAATAGGTGCACTCATTAAA  
AACACTACGGAAAAAACACTGTATTTGCCCGCGT

SEQ ID NO: 1060 ACATACTGCTGAATTTAACTCAAAATATTTTCAGGTAAGTGAAAGTGGTGCTTA  
ATGTAGACTATAGAATGACTTTTCAGGTGTTTTCAACTGAAAGTATATATCCAGAACTGCATCCTTA  
TAGAAATACAAGTAAGACTTAGGATAATTTGCCTTCAAAACAGTTTTCCTAATCTCAGCAGTATCC  
AGTGAGTGAAGAACACTTGACTGACTCTNNGGCCACCTCTGTACTTACTGT

SEQ ID NO: 1061 ACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCCTGCACGACCC  
ACAGAAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACATGCTTTCTCTTTTATT  
CTGTCTTCTTTTFIGATCTTCTTAAATCTGATTTTCTCCTTTTTCCCTCTCTTGTCTCCATCTTTTTT  
CCAGTCTCTCGCTTCTCTTTATTTCTCCCCATATTTCACTCTGTTCTCTCACCCCAACCTTCCCTTC  
TGTTTCTCTCCCTCCCTCTCTTCCCACCCCTGCCTGGCCTTCCATATATCAAGCAGAGTTTATCA  
CCTTATGCAGGGGCAGCCCTGCCACCTGCCATAAAGTTGATAGGCTAATGGCATTGTTGGATATT  
GCCATGTCACAAAGTCCAGGACAGCATCAAAAATAGCCCTGATGTCTAAACCACTCAGCTATCTTT  
TTTATTTTTTAAAATAAATACATTACATGCTTTTAAAGAACTATAAAAATATATAAAGTAAAAAGA  
TCTTTCTCTCACTGCCTCCACCTCTCCTGGTCTACCGTTGTGCTTAGGGGAAACCATTTGTGATTA  
GTTTCTCCTGTGTCCTTCCAGAGTGTCTTTATGCAAATGAAAATTATTGNGATAATATAT

SEQ ID NO: 1062 ACAGACCAGTGAGTCTGGGGAATTGCGGTCTCCACCAAGATCTGTGGGTGCA  
CTTGGCATGTTTGCTGCAGAAAAGGCCCCAGAATGGGCTGGCTTGAAGTGGAAAAACACACTTTC  
TCATCCCTTTTGGACCACGAGCTTCTTGAGAGCAAAGCATGTGTTTGATATTCCTTTGCTCACCCCTC  
AGGCCTTGTTTGGCAAATTGCCTGGGATACAGAAAAATAAGGACAAGGTCTGGGTGTAGTGGCTTA  
TGCTGTAAATCCCAGCACTTTGGGTGACCAAGGCAGGAGGATCTCTTGAGGCCAGGAGTTGCAGA  
CCAGCCTGGGTAACATAGTGAGACCTTGTCTCTGCAACAAAATTTAAAAATTAGCCAGACTTGGTG  
GTTCCCACTTGCAATCCCAGCTATTTGGGAGGCTGAGGCGAGAGGATCACTTGAGCGCAGGAATT  
TAAGGCTGCTGTGAGCTATGATTGTGCCACTGCACTCCAGCCTGGGTAACAGTGAGAGGCCTCATT  
TCAACAATAAAACCCAGCTTGGGCCGCGCGGTGGCTCATGCCTTTAATCCCAGCACTTTGGGAG  
GCCAAGACGGGCAGATCACGAGGTCAGGANATAGAAACCATCCTGGTTAACACGGTGAAACCCCT  
GTCTNTTCCAAA

SEQ ID NO: 1063 ACNCTAAACAGTGGATTGAGTTCCANCGNTTATTCTTTTTTNTCTTTTTTTCANA  
TCACCATCTAAGTTACATCTTTAGCTCAGGTCCATCCTTCCGAGATCTNCTTCTTAGCCCCCAGCC  
CCTGGTGTCTGTGTGGTCAAGTGACCTTACTCAGGAGCAGATATNTCCTTGGCCGCCATGGAGCC  
TCATCCATCCACACGTGCCTGTAGCATTCCAGAGCTCACTGNCCTTCTAGATGTGCCTTCCCGCTTG  
GCTTCCAACGGCTTGTGCTCACTCTGTCTGCCAGGTATGAGAAGAACACGTAAGACCGCCACCAC  
ACTCACCTCCCTCAAGGCCCTTGTGCCATAGGGGTGGCCACCCGACCTGCCCCANAACTTTTGG  
ATACTGGANGCAATTGCATA



SEQ ID NO: 1071 ACTGCAACTGCCAGAACTGGTATTGTAGCTGCTGCCCGCTGACTAGCAGCTG  
GACTGATTTTGAATAAAAAATGAAAGCATTAAAGGGTTCCCTACAAAACATTTTTCTTTAAAATAC  
TTTTGAAATGGCTATAAGCAGTTGACTTTCACCCTTGGAGAGCATCACACTGTGTGAGGTTCACTG  
ATTGTTGACCCTCCCCAGCCCCTCCTGCTTCTTTAAGTTATCTGTGTGCGGTGCGCTTCCTCTCAATCT  
TCTTTCACGCTCATTTCTTTTCTCTGACCCATGAGAAAGGAAAAGTTACTGATGATAATTTTAA  
ATAGTGTAATTATTCATTATAGCATGTACAGGATAAAATAAAAAGAACATTTTGTCTGGAAATGCT  
GCCCGGGAGCCTATTGTGTAAAATGTAGGGATTTTGNAANAATAACCTTGAAATTGNAATTGAC



ACGTNGTTNGGTCAAATTGNGTCAAAGTTTAATTTGGTTNTGTTTCCTTGNTTCCTTATTTTAAAG

SEQ ID NO: 1072 ACAATTGTCTTTATTATCATATATGGTATGTATGTATGATTTCTTTCCATTCCCT  
CATATTTAGACTGTATATTTATGTAGGTGTGAGTGATTGCCTGGTGCTTGCTTGCCAAAGGTGCTA  
GGCACCCCTCCAACCCCTGCCAACTTTTGTGGCCTCCCAAAGCATTCTGTATACCAAAGAGGCTTCAA  
ACCTGACCCCTCACTTCTCAGTGGACCCGAGTTTCCCTCCCATGCCATTATTTTCAGTGGGGAAGTTT  
TAGAGGTGAGCTGTTGGCCACAATATCAATTTTAAAGTGTTCATAGCAGTTATGTCTCCTGCATTCTT  
GGCTCCTGGATTTACCACCAAGAGTCCCAAATATTAATGCTCTTCCCTTTTCTACCCCTCAAACCTT  
ATAGTTGTATCTTATTTTTTAAAATGAATTTTCATGGCCAGGCACAAGTGGCTACCCGCCTGTAATC  
CCANCACTTTGGGAGGCTGANGCANGANAATTGCTTGAACCAGAAGTTTGAGACCAGCCTGGGCA  
ATGTANTGAGACCCCCTGTCTCTAAAAAAGGTTTGAACCAAGTTTGAGACCAGCCTGGGCA

SEQ ID NO: 1073 ACAAGAAGCTCAGCGCAAAGCGTGCGGATTTGCAGTCCACCTTCTCTGGAGG  
ACGAATTCCAAAGAAGTTTGGCCGAGAGGCACCAGCCTCAAAGAACGGCTGTGCCAAGATGGCT  
TTTACTACGGAGGGGTTTCTTCTGCCTCGTATGCAGCTTCAATTGCAGCAGCTGTGGCTCCTAAGA  
AGAAGATTCAAACCACTCTGAGTAATCTGGTTGTTAAGGGCACAACTTGATCATCCAGGAAACA  
CGGCAAAAACCTCGGAATACCCCAGAAGAGCCTGTCTTGCTCTGAGGAGTTCAAGGAACTGATGGA  
CCTGCCGACGTGTGGAGCCAGGAACTTAAAAACAACATTTAGCCAAAGCCACAGCTTCAGGGATTA  
TGGGGAGCCCAAAACCAGCCATCAAGTCCATCTCGGCCTCAGCACTCTTGAAGCAACAGAGATCAG  
CGGATGTTGGAGATGAGGAGAAAGGAAATCAGAAAGAAATACAGAAGCGATTTCTGCAGAGCTCAA  
GTGAAGTTGAGAGCCAGCTGTGCCATCTTTCATCAAGACAGCCCCCTGCTCACCTCCACGGACAG  
GATCCGAGTTCCCCANGCTGGAGGGAGCCCCGGCCACAATGACGCCCAAGCTGGGGCGAGGTGTC  
TTGGGAANGAAATGAT

SEQ ID NO: 1074 ACAAATCTGGAGTGGCTGTAAAATTCGGTTCTCAGAGATGAACTTGCAGATT  
CGGACTTTCAATTGTTCTGTGTTTGTAGTTTCTTATCAACTGGGGAAGTGTGTGACTAAGCTTT  
GTTAAAAGTAGAGAAGAGCTTTTCATAGTTCCAACATTAGTTGTTACCGGGCGCAGTGGTGTGTGT  
CTGTAATCCCACTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCCGGGAGGTGGAGGTTA  
CTGTGAGCCGAGATCGCGCCATTGCAACCCGAGCCTGGGCAACAGAGTAAGACTCATCAGCTCCCA  
ATCCCCAGATTCTCCAGTACAACTCTTCATAGTTTACTTGACCATCACCATCAATATCTGCTTCCC  
TGATCATTTTCATCAACTTCTTCATCTGTTAACTTCTCCAAGGTTTGTGCATCACATGGCGAAGTCTG  
CAGCACTAATATAGCCATTGCCATCCTTATCAAACACACGGAATGCTTCTTAATTCTTCTTCACTG  
NCTGNTCTTTCATTTTTCTTGCCATCATTGGCAAAAATTCAGGGAAGTCAATTGNGCCATTACCATC  
AGCATCTACTTCATTAATCATGTCTGTAACTCTGCTTCTGTGGGATTCTGCCCAANANATCTCATT  
C

SEQ ID NO: 1075 ACGCGGGGGGATGGGAGGGGTGTTTCATGATCATTGATATAGCAATCTACT  
CTGAGAAATGGAACACAGGGAGTTACCTATCACTTTCACTTATAATTCCAAAAGATGACTACAAC  
CATGTCCATGCTCAGATTCAAACAGTTTTCATATCACTTTTGGGTGGTAAGATGATTAAATTACA  
GTTTTTTTTTTAATTGGCAGCACCCTAACCATTTCCTTACATTCTTTTTTGTATGTGTGGTTTCTTT  
TTATTTAACCCTGAGCCGACATCGTAGTTTCTTGTTTTGTTTTGTTTTACAGAGCTGTTGCATGACT  
TATGTTACCATCCTAAAAAACACTATATTAACATGGAATAAATTGTCTTTTATGAATTAGGCTTT  
TTGAACATCCTGTGTGGGATTTTTTTGTTTTTCAATTGCAACAAAAGCTCTGTAGGGCTGACANACA  
TTTAAAGTTCACATAATCATCTGTAAACATTATGTAATTTTGTGGAAATACTANAATTTTTTCTCTG  
ATTTGCCATTATATGGATTGTCTATTTTTTGATTAATGCAAAAAGTATATGACTTTGNTTTTTATGTG  
ATAACCATAAATATTAAGGTGTTGAATACTAACANTGCT

SEQ ID NO: 1076 ACATGGAATCCTTTGAAGGTATATTCAAAGAACAGTATGATACCATCCATCG  
CTTGGAACAAACAAGTTGCGAAATGTTGCTAAGATGTTTGCTCACCTTTTATACACTGATTCACT  
TCCATGGAGTGTTCTTGAATGTATAAACTGAGTGAAGAAACCACTACATCATCCAGTAGAATTTT  
TGTCAAAATATTTTCCAGGAAGTGTGTGAATACATGGGTCTTCTAACTTAATGCAAGATTAAA  
GGATGAACTCTGCAGCCATTCTTTGAAGGATTATTACCCCGAGATAATCCAAGAAACACTCGGTT  
TGCCATCAACTTCTTACTTCTATAGGCTTGGAGGTTTAAACGGATGAACTGCGGGAGCATCTCAA  
AAATACACCAAAGGTCATTGTGGCGCANAAACCAGATGTTGAGCAAAAATAAATCCTCCCCATCCT  
CTTCTCTTCAGCGTCTCCTCTTCANAGTCTGACTCATCCGACTCTGATTCTGACAGCAGTGATAG  
CAGTTCAAAATCTTCCAGTGAAGAAAGCGACTCTTCATCCATCAGTAGTCATAGCTCTGCCTCAGC  
TAATGATGTAAAGAAAGAAGGGCATGGGAAGACCAGAAGTAAAGANGTAGATAAATTGATCAGA  
AACCANCAACA

SEQ ID NO: 1077 ACGCGGGATTTGACAAAGATGGTGATGGAACATAACAACAAAGGAATTGG  
GAACTGTAATGAGATCTTGGGCGAGAATCCACAGAAGCAGAGTTACAGGACATGATTAATGAA  
GTAGATGCTGATGGTAATGGCACAATTGACTTCCCTGAATTTCTGACAATGATGGCAAGAAAAAT

GAAAGACACAGACAGTGAAGAAGAAATTAGAGAAGCATTCCGTGTGTTTGATAAGGATGGCAAT  
GGCTATATTAGTGTGCTGCAGAACTTCGCCATGTGATGACAAACCTTGGAGAGAAGTTAACAGATGA  
AGAAGTTGATGAAATGATCAGGGAAGCAGATATTGATGGTGTGATGGTCAAGTAACTATGAAGAGT  
TTGT

SEQ ID NO: 1078 ACTGTCACTTAACCCCTATTAACATACGGTGTTCAAGCCTTCCAGTATCAGCG  
CCTTTTGCATAGCATCGCTGCTAGTCGTTTCACTAGCATGCTGGCTAGGAGTCTTCTTAGTCACTCT  
TTTCTTCTCCTTGATAGTGTCTGTACTATTTGTTTTTTGCTTTGGTTATTGTTTTTTCCCTTTTCTT  
AGCTATGAGCTGATCATGTCTCCTTCTCACCTCCTGCCATGATACTGTCAGTTACCTTAGTTAAACAA  
GCTGAATATTTAGTAGAAATGATGCTTCTGCTCAGGAATGGCCCACAAATCTGTAATTTGAAATTT  
AGCAGGAAATGACCTTTAATGACACTACATTTTCAGGAACTGAAATCATTAATAATTTTATTTGAAT  
ANTTAAAAAA

SEQ ID NO: 1079 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATGTCCCAGGCTGCCTTCCTGTC  
CTCCCATCCTGCATCCTGCTTTCTGTTCCCTGGATACCGTAGGATGGTTTTATTTCAAGTTCATGCAC  
AAATTANTCTGGACACTGNGGAGTCATAACAAGAGTGGGATGGAGGTTCCAGGGCCAATCAGTTT  
CTTTGGAGGGAATCTTGAAAAGNAGGTTAAAGCAATGCCACAAAAGCCACTGCTTCTCCTTCCC  
ATCCCCAGCAATGAGCTAANAGCCAACCTCA

SEQ ID NO: 1080 ACTGTGAAGTCAAAGGCCCAACATTACAGAGCGCACCTCTGCCTGAAATACA  
AAACTAAGTTAACTAGCAAGTTACAGGAAATAGCCTTCAGTAAATTCCACAAGCCAAGTGGCTAC  
TGCATTGTCCCTGAAGAAGGAGGGCCAGTGTTCTTTCTGGGTGTGTAAGGTCTTACTTAGTTCAA  
GGTTTGTCCTTCTTAGTTGTAGGTTGGGCCCTCTTTCGTTTCCAAGAGGTATAGGACACTAGAGTA  
CAATTGAATATATGTGCACAGTGATGAAACCACCACCACAATCAGGATAATAAAGAAGTCCCTTCA  
CCCCGAAAAGTTCTCTCCTGCCCCTGGATAATCCATCCCCTAATTCCCTATAGGCAACCACAGAA  
CTGCTTTCTGTCTCTGTTGATTAGTTTGCATTTTCTAGAATTTTATATAAATGGAATTATACAATAT  
GTTCTTTTTTAAATCTGGCTTCCTTTACTCAGCATGATTGTCTTGAGATTTATTCAAAATACAAGAT  
TTGTTCAAAATATGCTGAGTATAATGTGTTATCAATGCTCCATCGCATTTTGTGCTGAGTAGCATC  
TCAATGTATGGATGTAATTTGTTTATCCATTGCTGTTGGTGANGAAGCNCCGGGGCATTGGA  
TT

SEQ ID NO: 1081 ACCTTTGTGCATGTTGCCTTCATTCTCTGAGCAGGTATCATCCTCAGGGAACCA  
GCATGGCACCTACCAGGCCAGGCTCTGTTCTTAGGAGCAAGGAGCTTCTTGCGCTAACAGTTCTGG  
CCTGAGACCTGGATTGAGCCTTGGCAGACTTCTTGCTAAATGTTGGCCATTCAAGTCTCAGGCCCT  
CTGTTCCATGGAATTGGGAATCTCCAGGTGACCTAATCCTCATTGGTGGCTTGATGTTTGCTGGTAT  
CTTCCAAACTCAGTTCCAGACTAGATTGATACCTGGAGCCCAGCTGCCTACTCAGCATTTCCACT  
TGGGTGCTTCAATAGGCATTTCAAACCTGATGTGTTTAAACACTTGATTAGGCTCCGGTTTTCTTT  
GGCTTCTGCTTTTCAAGTGAATGGCATGACTGCCTATGTGGGTGGCAAGCCACCCANGTGCCGAGGA  
AAGAGACTGAGGGCACGAGCTGTTCCANTATAATAAATATATAAATA

SEQ ID NO: 1082 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTAAGGACCAGTATTGTATTTGACTTT  
TTAAAAAACCATTTTACCCTGGAGCTAAATACGAGGTTATACTGTTCTGCTTAATACAATACTG  
GCAGCCTAAAAATGCTTCAAGAAACCATATCCCCAACAGCGGCAGAGCATCGGGAGGAGACCCCTC  
TGTCTCTGAGGCTTCGGTGCATTTCTGCTCAAACGGTGGCGGGAGTGGAGGTCGCTGCTGAAGGGT  
GACGGTGTGGCCATGACACGGGCAGCAGCGGAACGGAAGACGCCGAGACCCAGAAGGCGCATG  
ACTGCCTGGCCTCGAGTCACTAAAGCAGTTTGATTTCACTCTTGCTTACTTTTCTAGTATGGCTT  
CCTCTACTCTCTTAACTTNTTCAGTTTCGACAACCGAGTTCCAGTATTTGAGATAGAGCCACATG  
AGTCTGCCATTCTGGCGTTTGTGTTGTTCCAAACATCACCACAGT

SEQ ID NO: 1083 ACTGTAAAAGTTCTGACACAAGACAGTGTTNGTGGTTACTTTTCATCGACTTT  
AGCATGTGATCTCAGGGACTCAGACATACGTCTAAGTTCTATTCTGAGTTTTGGCAACAGAGCAGT  
GACAGATATTTCTGAATGAACAATTTTTAGGTGTTTTAGCCATTTGAAAAGTATTGCCAACACAC  
TATTTGGTGTGTAGCTCAACAGTCACGTTGTGCCAAGAATTAAGAAGTCTAAAGTCTAGCAACATC  
TTACTTCACCAAGACTAACTATAATTGAAGGTTTACTATTTGTTTAATAAAAAATCACACATCAA  
CTTTTATCCAAAACAGCAACTACTACAAAAGGAATGACAAGAAAAAAATGACTTCACAGAAATA  
CACTAAAAAATCTGACAATGTTATGCAAGAACCCGCCCAAAGTTTNAAGTGTTAATAATGAATA  
GCACACNTGACCAANGTCCAAGATGTGAAGATNCCCATGTTCAAGAACTGGGGGGGAAATCACT  
CTACAACTAACT

SEQ ID NO: 1084 ACTGCTGCCGAAGTTGCCCCAGTTCCATGGGGTTCGTGTCTTTGGCATCAACA  
AATACTGAGGGATGGGTTTTGGGACAGCTCCATGGGCATGGGGAAGGCACTGAAACAGAGGACT  
ATAAAACATCCTTCTCTTATTCTCCATACTGTCTTCTACACCTTTAAAGCCTGAGAAGTATACAACC

TTTCCCAGACTCCCAAGAAGAGAAGAGATTGGCAAATGGGGCTCCTGGGCCCAGTCCTGCTAGTG  
GCAAGTTTCTTTGAATCANGAAGGCAGGTGAGGTTAGGGCCAAATCACTCTCCTCCATANCAGGA  
AGCCATTTGGGCAGCTTCTTTGGTGATNACATCTTTCCATATCTTTTACACTTTACCACCTTCCAGC  
TCTGTTTIGCTGNGTATTTTCTTACAATAATTTTTTTTTCAGCTATAGCTGCAGTIT

SEQ ID NO: 1085 ACAACCTTCAAACATTCCAGTTTTTTATAAAAAAAGGGGCACACAATCGTG  
TTTGATCCCCTTTTGTTTTTGGACAAATGTTTCTACAAATACAGATTCAACAAACCCAAAGGCTGC  
AAATTTACTTGCACAGTAATCTGCCAGCCCATTTACTCCACTTAATCCAGCTGAACTTGAAATGCN  
AACCAAATGTCCATGGTCATTAGCAATCATANCAGGTAANAAAGGCTTTATAAGTCCATAAATGT  
GCTTTGAAATTCACATCAAATGACTTTTCCATAAAGCTCATTTGGACAGTCAAGGAACCTTTTTGCC  
TGNTACGATTCCGGCA

SEQ ID NO: 1086 CGAGGTACAAATGTTCTGTAAAGTTGTAACAGAAAATGAACCCCAACTCTTT  
CAGTCTTTAAGACCTCAGTCACAGTCCCACACTAATAGCTGCCTTTTAGTAAAATAGTGACATCCA  
GTGGACAAATGAAATATTTCTCAAATTTCCCATATTTGAATCATAAACAAAAGTTGCCTCTTTAAAT  
CCAACAATTTTAGTATGAGTNCCTACTTATTNTGATGGAATCTACAGATTCTCATGAATACTTCAC  
ATCATAATTTCTATGTATCTTAAATATATCCTCTTNTCTCCTATTTAAAATCAATTATTTANATTGA  
GCTACGTNAGAATGTCTACTCTTGAATCTAGTTGANGTTNANGGTTTACTTNTTCTTTTNTTGCTT  
CANTTACTTCCCACCANCCANANTACCTGGGTTTGGCTNTTNGGNATGGNACGNNTTAAATTTNTT  
ANNNGGAANNANCTGNNNNNTCTTAATTATTTNATNTNTTNTGTTTTGAAANGNNTTCTCCANT  
ATNTGNAACTTCNANTNTGNTTGNNTTNTTTAATNTTTTATCNCNTNTNTGNTNNTCNNTNT  
NTTATGNNCGTNTTNTTNTTGANTCTGGTTTCTCCTCNTGNTTNGTNATCTTNTGTNATCTCGG

SEQ ID NO: 1087 GGTACTNCANCTANNGGCTCATTGGNATGCTATCNTGAANNACNTGGTNATN  
CTATATAANNNCNAATNNGNAAATCTNATAGCNNNNAANGANAAAATGNTGATTGNATTAGACC  
AAGTTAAGCAACAATAATGCATGAAACCAGTCGAATCATNAGCAGATAATAAACTGGATATCAA  
CTTAGAAAGGAGCAGAGTANCAGATATGTTTACAGATCAAGAAAAGCAACTTATGGAAACAATA  
CAGAATTTACAAAAAAGGATACTCAAACCAAAAGTATTATTTACAGAGACCAGTANTAATAATTGAC  
GCTGAAATTGCTTCCTTAAAAACACTGATGGAATCTAACAACTTGANACATTGNTATCTTTGC

SEQ ID NO: 1088 ACTTCTGTGAGATTACGGNCGCTATGACATGGCTCAGCTTCGGTTTTAAAAA  
GGTCAGTGTCTATCTGGAAATNTCTATGTGANAGGTGATGGAACCAGAGTTTACTTCTTCACACAA  
GAGGAACTGGACACGCTTTTCACTGCTGGACTGGAAAAAGTTNAAAACCTTGGTGGACCGCC  
GACTGCAGTTCCGNCANGCACGGNGGCTCACACCTGTAATCCCAGCACTTTGNGAGGCCCGAGGC  
GGGCAGATCACCTGANGTCTGGANTTCGAAGACCAGCGNGGNCAACNTGGTGAAACCCNTCTTT  
ACTAAAGATAANAAANAATCGGCCTGGGTGTTGTTGGNGNGNTGCCTGNTNNTCCAANTCTCG  
NGGAAGGCTGNGGGCANANAGAAATTACNTTGAACCCNTGGAGGCANGGTTGCCAGTNGA  
TGCTTNAAGATTTGCNCCACTGGNACCCTGCCCCGNGCNGGNCNTTCTNAAAAGGGGCTNAATT  
TCCATCCCCACTGGGCGTTCNNTTANCTTATANGGATNCCCCGAACTTNTGTNNNCNCAANCN  
TNGGCTGNNCNTCAATGGGNNCANTTNACNTNGNCTTCCCTNNGCGTNGANAATNNTTAATT

SEQ ID NO: 1089 ACTAAAGGAGATAAGTGTAAGTTTTCCCATGACTTGACTCTGGAGAGAAAAAT  
GTGAAAAGCGAAGTGTTTACATTGATGCAAGAGATGAAGAACTTGAAAAAGATNCTATGGATAAT  
TGGGATGAGAAAAAGCTGGAAAGAGTATNGAACAANAAGCACNGTGAGGCGGAAAAATAAAAAA  
CCCCAAAACCTCAAAATNGTGTGNAANNNTTCTGAAGCTATTNAAANCAACCAAGGTATGGGTTG  
NTTTTNGGGTATNCCCTGGANGGGGTNAATTTTGCATGNTCNANNTGCCNTTCCNCTGGNTTT  
NNGTTNAANAATAAAATAAAATANAANTGCTCTGCCCGGNGCGNCCNTTNNAAANGGNTAA  
ATTTCCATCTNCNTNGNNGCCCTN

SEQ ID NO: 1090 ACAACCCTGTGACTGTGATTGGGTTATTCAACAGCGTAATTCAGATTCATCTC  
CTCCTGATAATGAACAAGGCCTCCCCANAGTATGAANANAACATGCCAGATNCCATAAGGCAGC  
CAAGCTCTTCCANGGGAANATTNTNTTATTCTGGNGGACNNNGGTATGAAAGAAAAATGGGAAG  
GGATTTCATTTTTCAAACCTAAGGAGTCTNAACTGCCAGCTTTGGCAATTTACCANACTCTAGATN  
ACGAGNGGGATACACTNCCCACAGCAAAAGTTCCCGTAGACATGTGCAAACTTTTNGATNGA  
TTCTTAAGTGGAATAATTGTTGAAAGAAAAATCCGTGAATCACGAAGNAAGACTNCNAANGTGGAAC  
TTTNAACCTTTCTTTGGACTACATATNGCCAANTTTTACCTTTTTTNNAAANTTAAAAANGCNCNAAT  
CATATTNTCAATGACACTTAACAACCTGGNTTACTTTGNCTTGGNANTCNNACACANNANNCNCTT  
GTACACAANCTCGGAGATCTTGNCATATANACNGCACNNCANNTAACNCAGGANNTTNTTTC  
TGTCTTAAANCTCNGTTTCNCNTTTTTCTNTTNTTAAAAATNTNNATCNTNCTCCCTNTNCNNTTNT  
TTTTNT

SEQ ID NO: 1091 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTAAAATCCATCTGACTGGG  
TTAATTCAAAAAACCTGTNTCCAGCTCTNAAATCTTTCTTCTCTTCTCTAAACTAGTNTTAAA  
GCTTNCAACTNTNTTTNGTAATCTTTCAATGAATTTTTTCATTTCCAAAANATGTTNNGGTTTATTT  
TTAGTGATATCTCCTTTNATATCCCAAACCTGTTTTCTGATGTGTGNGGGTTTNAACCTTCTCTTGN  
ATCTCATTGAGCTTCTTTAAAACCAACATTCTNAAATCTTTATCTTGNATTINAAAAATTTTAATTT  
TGGTANGANCCATCGCTGAAAAGTTAGNNGNANCCTTCTGGACCACTGTAACACACTTTTTTCATA  
TTTTNAAAAGTCCNNGTACCTNNGGCCGGAACCACTAAGGGCAAATTCNAGCNACTTGGCGGC  
CTGTNCTANTGGATCCAACCTNNGTCCAANCTTGGNGNAAATCTNGGCATANCNNTNTCCTNNGGGA  
AATTNGNTNCCNCAACAATTTCANACNNNTNTCNTACCNGAAANTTATANTTNTAAANCCCCGNGG  
GNCCNNATANNGANANCNNACTCNNATNATNGNNNNGCCTNNATGCCTTTTCAATTAGGAAANC  
C

SEQ ID NO: 1092 ACTTTTTTTTTTTTTTTTTTTTTTTTNGNTTNTTTTGGGGCAGTTCAAGTTTAATACAA  
ACTACAAAAGATTAATGGGTTGCTNTACTAATACATCATANAAACCANTANCNTGCCCANAAACNC  
CAACTNAGGCCNTTNTACCAAAGGAAGAAAGGNTGGTNTNTCCACCCCTNTAGGAAAGGCCTG  
CNTTGTAANANACCANANTTCGGTTNAAATNTGAAGTCTTGTNTTTTACTANTTGGAAAAAAAANT  
ACANAAAAGGTTTTNTTNTAATGGTTGCCNACCGNCANCTTGANCTAAAANANCCNAAACNTTA  
TTTTTTGTTTGGAAAAATATNTTNTTTTTTTGAAAAATAAGGCTTINATGGTTTAAACTGNCAAGT  
TTTNNANCNANNTGGGNNANACTTTCNNATNTTTNGNAANAACTTGAAAAGCNTAAAATTATT

SEQ ID NO: 1093 CATCCTTACAAAGATTTCTGCGNGTGATTTGTGTGAAGAAGAGAACGTTTGTCT  
TTTTCAATGAAGCTTTGCAGATCACCATGTGGTTGAAGGTTTTAGTTGTGGACACAGTGGTCCCTC  
CTTANTGATGAAGATCCTGCCTTGGGCTTCATGGAAAACATGCCACCCCTGGGGCTGCGTTTNGGA  
TTTATTGTTTTTATCCACACTTCTACTTGGTCTCTTGGATTTCTAANCANCATGCANCCCTATTGT  
TTNNTNNGTNTTGGAAANNCCNCCCANTAGAAGGNGTTTTTTTTTTTGGTCACNNCNCNNNATTT  
NATNGTANANNNTNCNTTNTTNTTCAAANGGNNNATCCCCNTNTTTTTTACCNTANNATAACTTTCN  
ANGNCTTATGCGCTGATTATNAAAA

SEQ ID NO: 1094 ACCTGGTNAANCACTGTGGCAACATACCTGTCTTCTNTTATTAATTATCCATTA  
NCACTCAAGCCTTTCTACATGAGGGATAATGAAGATGGCCCTCANCACACGGNTGCTGCTGNTGA  
TCTTCNGGTTCTGNAGNNGGGGAACCTTTGGANGANGCCTCAGATAAGAACGATACCATTCTT  
AGAGGAGCGCTTACCAGATCGGGACTTACANAAGTCTACCAATGGTATCNGGACCTTNGTCNATT  
NNGATCTGTNCCACATGGANGTTTNGGGATGGGATTNCAACGCTACCNGCNATGCATNNTGGGTG  
TTGACAATATCAAAAANNTNATCCCTNNCNNAAGGNTTCTNATTGATNCNTTTNNTAGCTGGNA  
NATTGGTTNANGAAAAGCACCCCCCATGGCNNAACACTGCACATGATTGNGCATACTANCAATA  
ATGCTGTTTGGATTNAGAACTGCAAATTTCAATATNTAATNGNNGTGCCNNTNGGATATCATANA  
AAANATTNGGTGGCGATTTTTNTTNGGAATTAGAGGGNTATNAAAAAT

SEQ ID NO: 1095 GGTACGCGGGGATTCTTCCCCCTCTCTACAACCCTCTCTCCTCAGCGCTTCTTCT  
TTCTTGGTTTGATCCTGACTGCTGTGATGGCGTGCCCTCTGGAGAAGGCCCTGGATGTGATGGTGT  
CCACCTTCCACAAGTACT

SEQ ID NO: 1096 ACTGCCCCTTTACATCAAAGAACTACTGACAACGAAGGCCGCGCCTGCCTT  
TCCCATCTGTCTATCTATCTGGCTGGCAGGGAAGGAAAGAACTTGCAATGTTGGTGAAGGAAGAAG  
TGGGGTGGAAGAAGTGGGGTGGGACGACAGTGAAATCTAGAGTAAAATCAAGCTGGCCCAAGGT  
GTCCTGCAGGCTGTAATGCAGTTTAATCAGAGTGCCATTTTTTTTTTTGTTCAAATGATTTTAATTA  
TTGGAATGCACAATTTTTTAAATATGCAAATAAAAAGTTTAAAAACTTA

SEQ ID NO: 1097 NGTACCACCATTGTGNACCTTAACGAAGAANAANATCTTCAAGTNGACCCTAN  
AAGAGAGTTTTAAAAAACAAANCAATGTAAGNAANGGATATTTCTGNATCTTNAATTCATCCCA  
TGTGTGATCATAAACTCATANNAATNATTTTAAAGATGTCCGGAAGGATACTTTGATTAANTAAA  
AACACTCATGGATATGNTAAAACTGTCAAAGATTAAAAATTTANTACTTTTATTTATNGTTATTTTA  
TTGNAAGAAATAGTGATGAACAAAGANTCTTTTTTCACTGATACCTGGTGTATATTAATTTGAT  
GCCATCAGTTTTCTGAAATGGATATTTCANATTTGCTTNAAGAAAAATTAATCATCTATCTGAGT  
AGTNAAAATTCNAGTTNAGGNNAGCCATATNNACAACNTTTNNGGCNAATGTTTATGTTTCNNG  
TATATGNTCCCTNNCCGGCCGCGGNTTAAAGNCGAATTCNCNCTCNCCTTGGGCGNCCGGANTTA  
GGGGANTCCGAACCTTGGGCCCAANCTTGGNGTAATNNTGGCTNAANCTGGGTCCCNNGGGAAAT  
TGNGNTCTNGCNCNNAANTCCCANAAACATTTGACCGGGAAGCTANAAGGNAAANCCNNGGGGNC  
TAANGTNGGGGCCNCCCANTAATTGGNGGGCGCCCATGNCCGTTTCAAGTGGGANCCNCGNCACT  
NTG

SEQ ID NO: 1098 GGTACGCGGGATGGACTCTGCCACTGCCCCGACAAGATCAGAAAAGCTGTAT  
CTCTATGCGGCTCATGATGNGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAA  
TGGCCNCCGTTTGTGNTGACCTGACCATGGANCTTACCAGCACCTGGAATCTAAGGAGTGGTTT  
GTGCAGCTCTATTACCACGGGAAGGAGCAGGTGCCNAGAGGTTGCCCTGATGGGCTCTGCCCCGCT  
GGACATGTTCTTGAATGCCATGTGAGGTTATACCTTAAGCCCCNNAANAATACCATGCACTCTGCT  
CTCAAACCTCACGTGATGGAAGGTNGGGAAATGAANAGNNACTGATTTATAAAAGCAGGATGNGT  
NGATTATAAAANAAAAGTGCCTTTTTTTACCNATNGGGNTAAANCNANNTNTNNNGNAAANGNT  
NTCCTCTCTTATGTTNGTCTTNCNACNCNTNNGTNCNCTNCTTAATCNTGAATNTGNCNNNACTT  
NTCCATCAANTAATAGNTTCNTNCCNGGNNCGGNTNGTTTTNAANGGGTATAATTTANNTCTCC  
CTNGGGGNGGGCGG

SEQ ID NO: 1099 ACTTGTNCCTAGTTTTTCAAGGTATTGGCTGTTCTATAGATGCANTGATTGTC  
CCAGCTAGCTCTGTTACAGTCTTTTGGTGTGNTTTATGTNCAITTTGGAGAGTCAGGGCNAAGA  
CAGGTGATGTAGCACTNCTGTTTTTAATAATTATTGCTTAAAATACCTATTAATANTNTGGGTCAT  
TTAAAGGGACTTGAGGAAGCTACCCAGGATTACANAAGAGNGTCCACCTAACAAGATGGTCTGGC  
NGNNTCCTAGTTTTGTNTCTGGTTCAATAGAAATATGTGAAAGTGGTAANGTCATCATTGTATGCA  
GAGANCNGGTTACTCTATAATAAATCCCTTTGCCNAATGCATGAGAATGCAGACTTGCTACTGGCA  
NNNAGTGAAGCCCGTGGGTGACATAAAACTATTNTGACGGNGNNAACCGAATTCAAGANACGG  
ANTTTTATTCTNTTTTCANNAAGAGTCCNGGGCAANAAATTTGTACCTTCNNNGGAAGGGAACNN  
NGNATTCTCCTGGANGGGCACCTTGGCCTTTCCATTNTGGGGNNGANGTTTGNCTNTTTCTGTT  
GGCNNNCCAGGATGCCNTNCCCTTNNAAAATATTNGGGGGANTATACCNATCCCCCNNTGGGG  
GANGGTTTATGNTTTAANCACAAAANGGTTTCNTTGGNAAACNGANNACTTTTTTNGGGAAGNAA  
AAANCTCAAAAAATNTTTTGTGNA

SEQ ID NO: 1100 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGCTNNTNTTTTTTTTTTTTTTTGA  
ANNNNNAAAANCAAAACAANTTTATNTTAAAAAATACCAAGNATCACATNCATANCCNATTCCTAA  
TTNNANTTACTTCTTAAANCAACNTTTGNNTTGTGTTACAAANNNGCNCNACAGNGATAAAACCTA  
ATNTCTNAAACTGATTGNTTACANTTTATAANACAAACNTNTNTATATGTCNAAATCACATTTA  
GGCAGTTTAACTTAAATGCATAGACAAATTNCATNCNTTAAATTTGTCATTACTGNCAGNGAAAA  
ATCACAATTTCAANAAAAAATTCNACAAACTCATAGANCACANATAANAAGTCAAGTTNACTTA  
ANGGCAANAAGTNATTATCTGNTGATCANTGACAGGCATTGAAANTGCCTTTATCCAGGGGTAAA  
NTGAAAAGGCAANCTATCCTGTTACTTTGCACACTTGAATTGTTGCGCTGATCAACACTACNATNT  
TGNAAAACATTAAATTTCTTAAAAAANGNTGCCACATGTCCCCNTAANATGCCCAAN  
GCTGGTACAATTTCCACCAANATNGGTCCTAACACGTTNAGGNAAAACCTAANAATGAANGGGNTT  
TGCAACCAAATC

SEQ ID NO: 1101 GTTCNCGGGGAGAANCTTGGACCGCATNCTAGCCGNCGACTCNCACAAGGCA  
GAGTTGCCATGGNGAANANTNCAGTGGNNNCATTNTTGCTCCTTGTTGGCCCTTTCTNCACTCATG  
GCCAGAGANACCACAGNCANACCTGGAGCCATGAAGGACACAAAGGACTNTTGACCNAAACTGN  
CCCANACCTCTTCAGAGGTTGGGGTGACCAACNCATTTGGACTCACACATATGAANAATCTCTAT  
ATNNANTTCNAGACANGNAACNAACCCTGGATGATTATTCATCACTTGNATGANNGCTCTANACA  
GNTNAANCTTTANANAAAAGAGNNTGCCTGATAAAAAAATTAAATTCAGANAATTGGCNATNCNT  
GNAGGGGCCTTCTCAATACTTGGNTNATGAANACNAATTGNNCNANCAACCTTNTTCNGANGGC  
CANNTTNTCCCCTGGANTNATNTGTGTTGGACCTCCNNNGNNTGTTANTAGCCCGGTTNNANTTG  
GAAGANTTTCTAACCCGTNTTTATGNTTTCGTANCCNNGNATATCCANNTTTTGNNTNGTTGACAA  
NNANTATAAAANCTNG

SEQ ID NO: 1102 ACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCCTTAATTGGGGG  
GAGGATNTNTTCCAACCTTGGGGGAAGGGTNAAGGCNTTTTACCTNTNAACCTCAATCATNACNTT  
NTAAAAAGTTNTGCAGNNCCCCNCNTTGNACAGGACAGGATGGAAAANTGTNCTGGACTGGCTG  
NATTTCCCANTGTNAANCTGGCANCNCNCCTNAGGTTNTTGAGNGTTTATCAACTCACCAGGTCAA  
AANTNTGCAAAACNCTGGGNGGANCAACCTTAACCCACAAAACCTAAGTTGTTNTTCCACTGGNGC  
ANCCCTGCCNTTGGCGTTTTNTCAAGGGAAANNCCTTGGGGTAAACAATGGTAAANATAGGCTC  
ACTCACCATAACAGTGCCCCAAAGGGCAANNCANCTTTAGGAACTTNTGGGTAATGAAAT  
TGGAAACAAAANGTTCAAATCTTTCANCAACCTGNGGNAAGGNNTNANACANCCNTNTTATACCC  
AATTCCTGCAAGGNAAGGACGTNATCATGCTCCGNCCTTTGGCNTTNACCTNAACCTGGGNAGG  
NAAANCGGGTTCNTGCCTTNATAAAAAANNNAAANACCT

SEQ ID NO: 1103 ACACTTGAAACCAAATTTCTAAAACCTGTTTTTCTTAAAAAATAGTTGTTGTA  
ACATTAAACCATAACCTAATCAGTGTGTTNACTATGCTTCCACACTACCCAGTCTTCTCACACTTCT  
TCTGGTTTCNAAGTCTCAAGGCCNTGACAGANAAGAAAGGGCTTNGAAAATTTCTTTACAATTCA  
NGTCTTTNANGCAACNTTGAAANCCCTTCTTTNATNTTGNCAAGCAACCAGANCCCTNTATTNTGN

NAGGNTCCGTAANCATAAAAAACCGGGTTTGAATTNTCNTTCCAANGANTTTCGGCTTTNACTNN  
GCAAAAAAATGGTNANNAANATTATTCNTGGGGNANTNCNTGCCNTTTNAGAAAAAANGGTNTT  
NNCAAANTTTNTTNTTAACCCAATTTNNGGAGGGNAAAAACNGTGNNTNNAANTTTTTGGANT  
TTCTTTTTTTT

SEQ ID NO: 1104 ACTGTTTATTAACCAACCAGCTTAGAAAAATAATCATGGTAGACACCTTAGTT  
CATTCTTCTAATAAGCCTGTTGATCTGGTCCCTGTTGCCAGCATCTCCACCTTCTACAAAATGG  
GTGGTCTTTTTCTTCATTCCACCTCGTGGAGAAGACAATTTGAAGGGCCACAGGAAGTTATTTGCC  
TCTTTGAAGCGTTTTCCAACAGTATAGATCTCATGAATCAAATCCTCCATGCANATGATGCCGTAT  
TTACCAANAGATCTANCAATCAAAGCGTTATCTGTCAAANCAATTTNCTTCTATTGATTTTGCCAT  
NACCANGCTTGNAGATTAAGNTCATTNNNTGACTTCAGATTGGGGTACCCTTGNTCNTNGANCACN  
GCTAATGGGCNAATTTTCATCTATNATTTNCNTGNCNGTTNCTTANCNNGANCCCNANCTTGGTNT  
CTNATCTTTGGTGNATANTNTNTGNATTANCTGTTTCCTCGTGTNAAAATTGTTANTNNGTTANA  
ATTCATAANNNTNATANCNANCG

SEQ ID NO: 1105 ACTTTTTTTTTTTTTTTTTGTTTTTTTTTCTCAGCTAAGCCCATACAGAAGAACC  
AGACTGATGAGGAAAAAGAACAGGAAGAAATAAAAGGAAAAATGAAATAAGTAAGTTGAGAAA  
GAATCCTGGAAAAAAGTAAAAAATNCAAAATAAAAAGGAAGGCAGAAGGAATAGGCNCTCAA  
GGAAGGAAAAAAGAANGAAAAATNTAGGCTATTTGGATAGTTTAGTTAAGATATACTGAAACCT  
TGTGCCAACCTGNANAGAAGATCAGCCAATAAAAGTCAAGCCTCCCTTACCTTTACCTTACCTTC  
CTGCCATCCCANTGTCACCTGGGTCTGGTCCCTATTATTTCATAACAATTTGCATTGTGTGTTCC  
TAGAATTTTCAAAGCTGGTCCCTTATGNANATTTGCTNTTTTTTTTTTNCNTTNCATTNTTCCCTTNTAN  
NTAATNCTNNCTTTTGGGCTTNAANGNNTNTGTCCTTTAACCANANGNNTTNCNCAACNTGNAAT  
TTAANNGAATTGCTTNNNCCTTTT

SEQ ID NO: 1106 ACTAAATGGTATCCTTAGATTA AAAATTTTGTGCTTGATAACAGCTGNTTTTTTC  
TACATTAGAAATAAGATGCCACACAAGGAACCTACATTCCAGATTTAAAGAAATGAAAGGATACCA  
TTAGTGTGTATAACAGATTATTGTTTCATACTTGTAAGCATCTTATGTCATTGAGAATATAAAGAA  
CAGTGCCTTAGAAGACAGTGAAAGGTAAAGCTCAGCTTAATGTCTATGATTTGTTCTTTGACATTA  
AGGAAGGTAAAGGATTGGTCAGAGGATGTAACTTGATGTGAGCAGTAGTAAACCTGTTTTAGATAT  
CATACTGNTAATATTTTATTGAAAATTTATTTTCAGAGCGGAGAACTTAAGCTAAAGTCTGTTATA  
CAGAATTGAAAGCCTTCGTATCTTGGACCTTCCAACCATTTTTTCTTATGGCTGTNGAAAAGTATN  
NAAGCTAAATNGNTTTAATACCACTTCCCTTGTACCTTTGGCCGCGAACACGC

SEQ ID NO: 1107 ACTTCGTGTGCTCCGACCCATGGTGACGATGACACACCCCTGGTGGCATGCCC  
GTGTATGTTGGTTTAGCGTTGCCTGCATTGTTCTAGAGTGAACAGGTGTCAGGCTGTCAGTGTTC  
ACACAAATTTTAAATAAGAAACATTTACCAAGGAGCATCTTTGGACTCTCTGTTTTTAAACCTT  
CTGAACCATGACTTGGAGCCGGCAGAGTAGGCTGTGGCTGTGGACTTCAGCACAACCATCAACAT  
TGCTGTTCAAAGAAATTACAGTTTACGTCCATTCCAAGTTGTAAATGCTAGTCTTTTTTTTTTTTTT  
TCCAATAAAAAGACCATTAACTNAAAAAATAAATNAAAAAAGTACCTNGGCGTNACCAC  
GC

SEQ ID NO: 1108 ACCAGTAAGGCTGGATCTTACAGAGAAAGACTATGAAATACTTTTCAAATCT  
ATTAATGGAATCCCTTTCCCTGGAGGAAGTGTTGACCTCAGACGCTCAGATTATGCTAAAGTGCC  
AAAATATTTTATAACTTGCCATACAGAGTTTTGATGATGGAGACTATTTTCTGTGTGGGGCACA  
TGCCCTCCGATTGTAAGAGCTTTCAGTCTGATTAGTGGAGAGTGCTTATTAAGTCCACAGATAC  
TGTTGACGTGGCAATGCCGCTGAACCTCACTGGAGGTCAATTGCACAGCAGAATGTTCCAGAATTT  
TCCTACTGAGTTGTTGCTGTCATTAGCAGTAGAACCTCTGACTGCCAATTTCCATAAGTGGAGCCT  
TTCCGTGAAGAATTTTACAATGAATGAAAANGTTTAAAGAAAGTTTTTAAATNGTCTTACTTTCCA  
AATTNCAGAAATGGGNAAGATTGGNNTTANTTTCACCATTTGGATGGATTTAAAGTATCCCNGTATAT  
GGNGTCCAAG

SEQ ID NO: 1109 ACGCGGGGATGCCAAGGTGATGAAGGATGCAAAGACGAAGAAGGTANTAGA  
GTCAAAGAAGAAGGCTGTTGAGAGACTGGAGGAACAGTTGATGAAGCTGGAAGTTCAAGCCACA  
GACCGAGAGGAAAATAAACAGATTGCCCTGGGAACCTCCAACTCAATTATCTGGACCCTAGGAT  
CACAGTGGCTTGGTGCAAGAAGTGGGGTGTCCCAATTGAGAAGATTTACAACAAAACCCAGCGGG  
AGAAGTTTGCCTGGGCCATTGACATGGCTGATGAAGACTATGAGTTTTANCCAGTCTCAAGAGGC  
AGAGTTCTGTGAAGAGGAACAGTGTGGTTTGGGAAAGATGGATNAACTGAGCCTCACTTGCCCTC  
GTCCTGGGGGATNAGAGGCATNNAGGCTTNNCANTNCCANCATCTTTTGGCAAAAGGNTAAACC  
TGGAAATNTTTTTAANGGNAACTTAATCAGTTNTCTATTGGNCTANCTTTTTAAAAATATTTNNN  
ATTTCAAATTTN